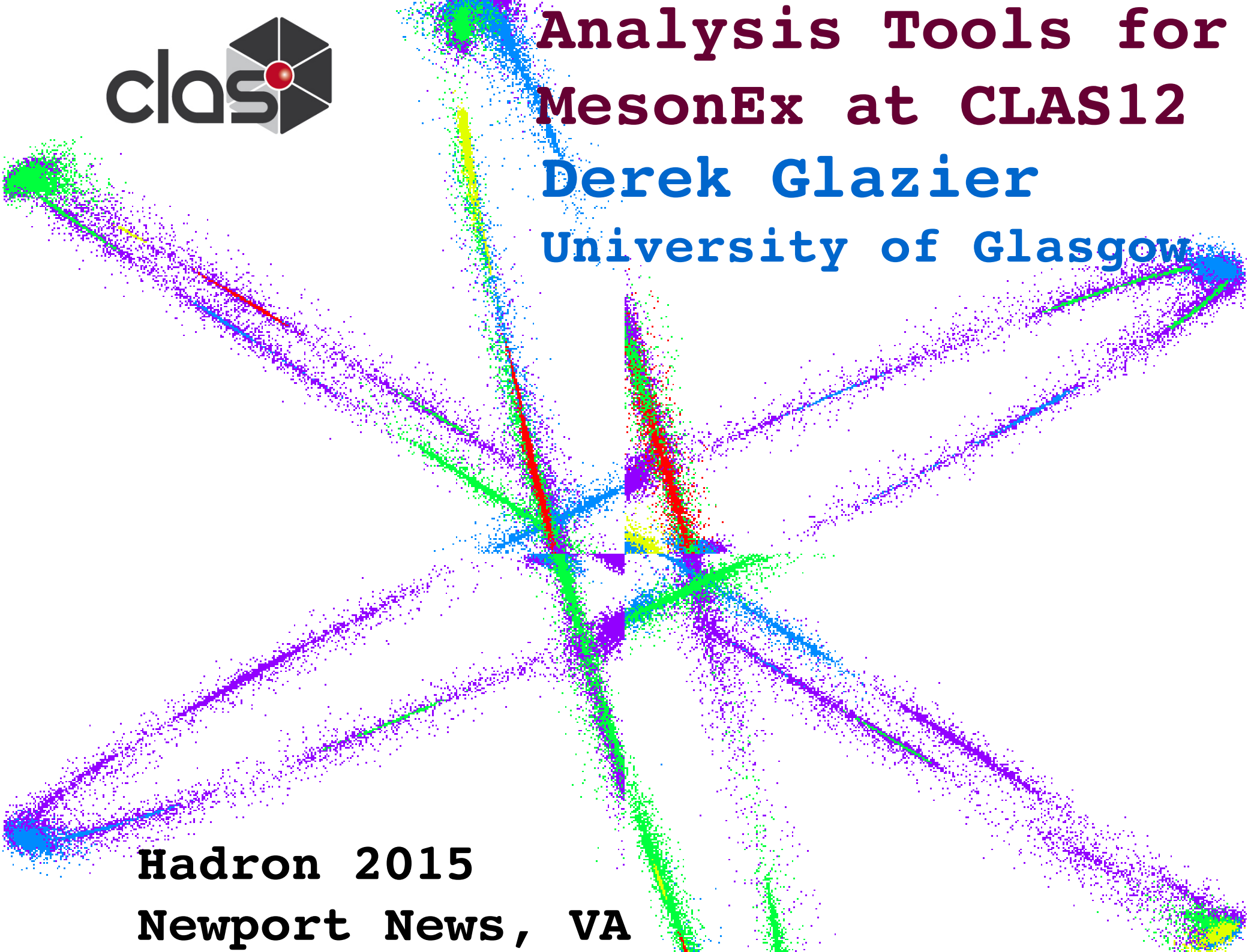
Analysis Tools for MesonEx at CLAS12

Derek Glazier
University of Glasgow



Hadron 2015
Newport News, VA

Overview

MesonEx

(also Carlos talk)

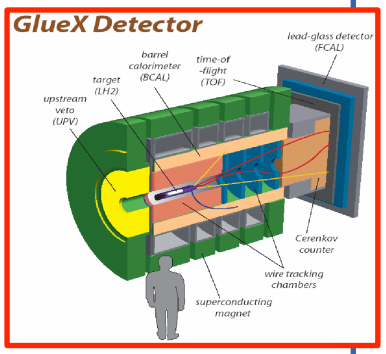
HASPECT

Data Handling Software

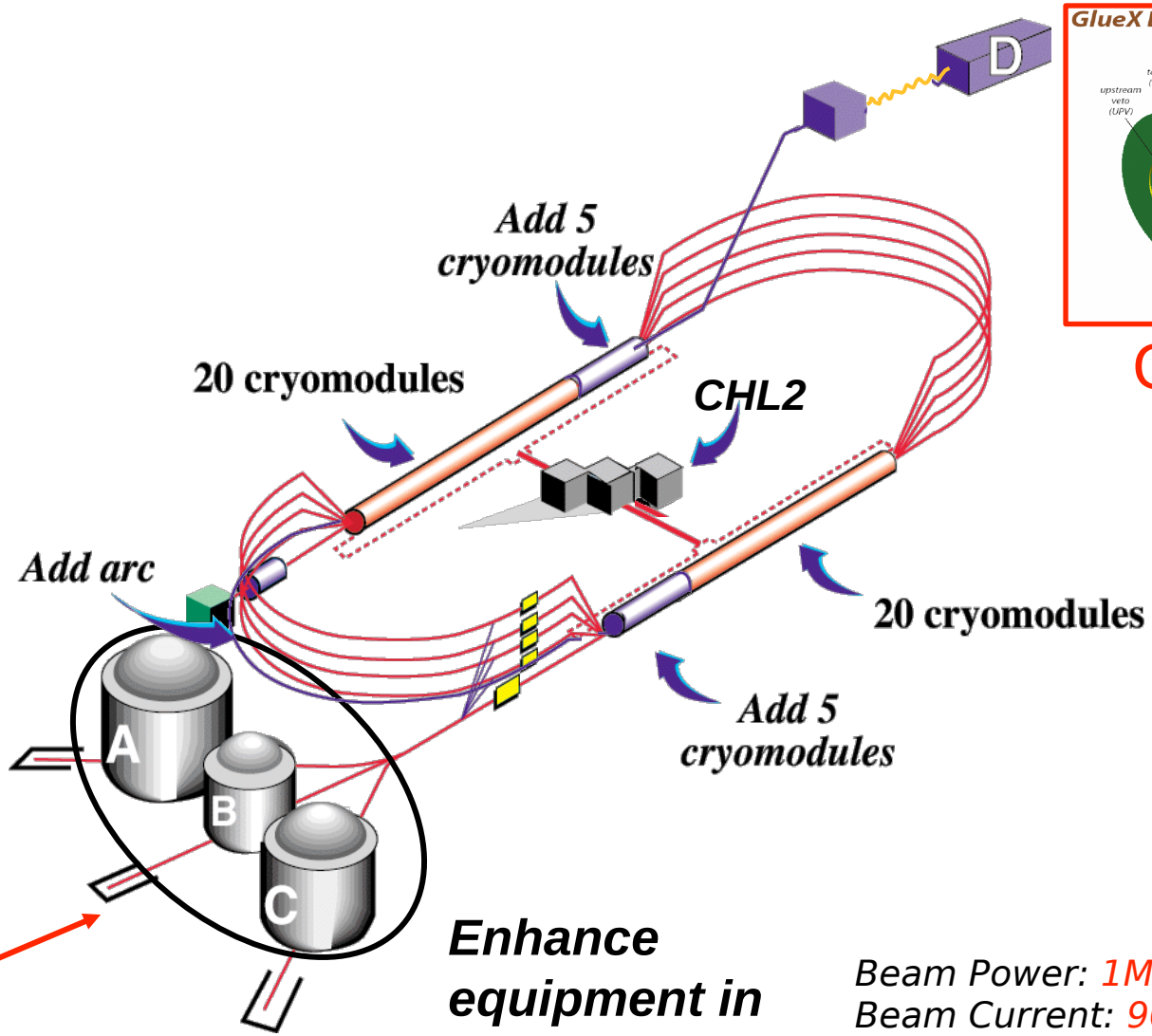
Longitudinal Plots

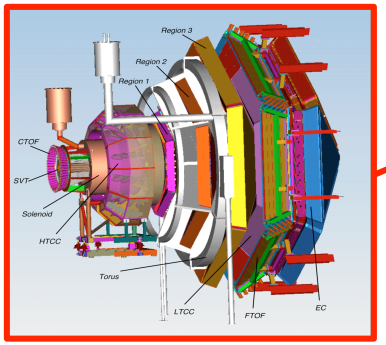
Likelihood Analysis

Jefferson Lab at 12 GeV



GLUEX





CLAS12

Enhance equipment in existing halls

Beam Power: 1MW
Beam Current: 90 μA
Max Pass energy: 2.2 GeV
Max Enery Hall A-C: 10.9 GeV
Max Energy Hall D: 12 GeV

MesonEx - Spectroscopy with CLAS12

Primary Physics Goal for CLAS12 e- beam

- Nucleon Structure(not discussed here)
- But high potential for Meson spectro. (see also Hall D GLUEX)

- Strategy :
  - High Intensity electron Beam
    - Tag quasi-real photons
  - Large Acceptance Magnetic Spectrometer
  - Many final states
  - Linearly polarised photons
  - Amplitude analysis sensitive to small contributions
    - Close interplay of exp - theory

HASPECT

International collaboration preparing for MesonEx at CLAS12

Implement reliable amplitudes
- Revisit techniques from earlier efforts
- Accessible to all
- Common Tools e.g IU AMPTOOLS
- Work closely with JPAC

Frequent meetings and Workshops
- HASPECT weekly meetings with experimentalists and JPAC
- HASPECT weeks with guests from other projects
- ATHOS Amplitude Analysis Workshops
- ...

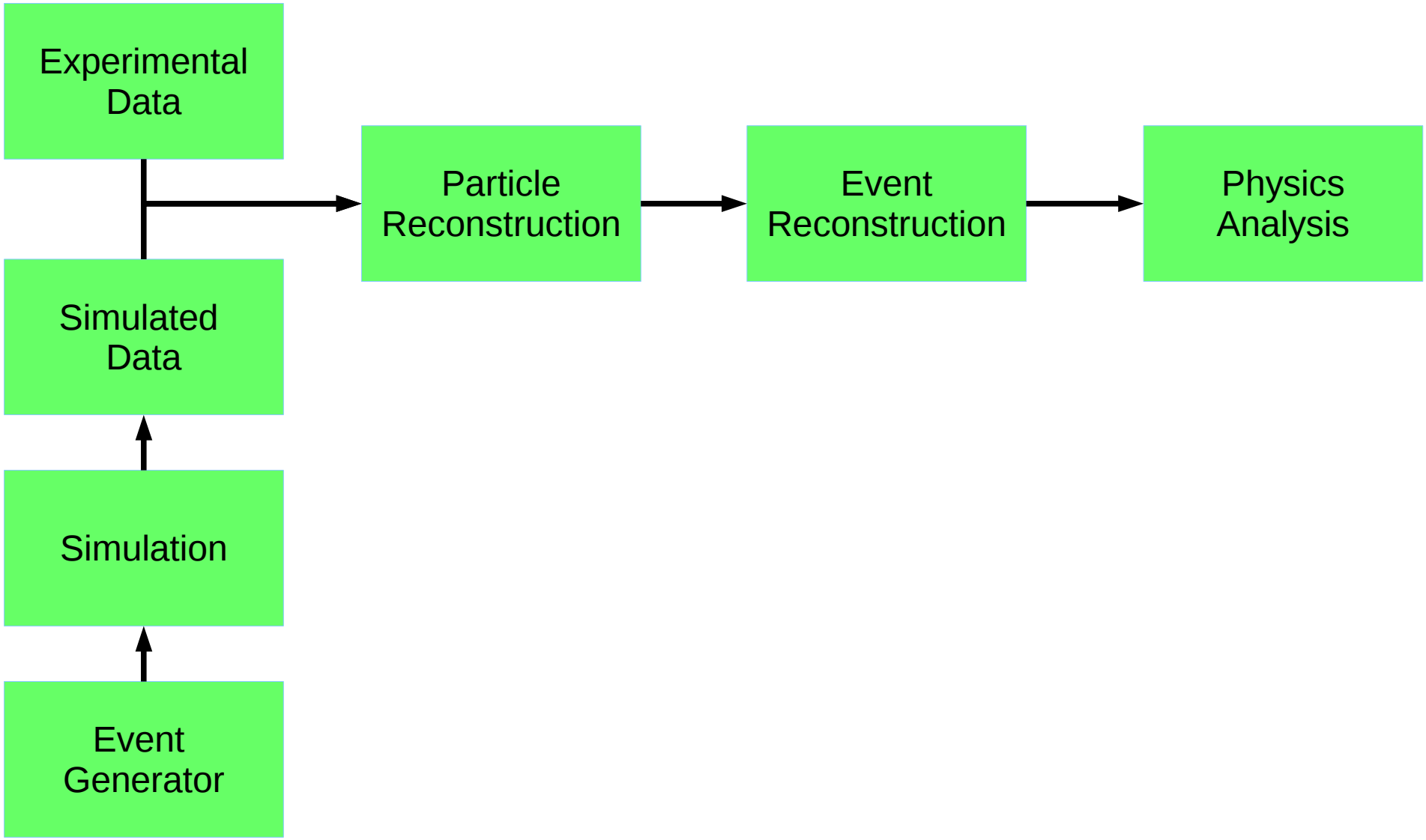
Apply/develop with existing CLAS data

$\gamma p \to N \pi \pi$ $\gamma p \to N K K$ $\gamma p \to N \eta$
$\pi\gamma p \to N \omega$ $\gamma p \to N \pi \pi \pi$ $\gamma p \to N \eta \pi \pi$
$\gamma p \to N \pi^+\pi^- \pi^-K^+$
$\gamma p \to N \varphi \pi$ $\gamma p \to N \varphi \eta$ $\gamma p \to N$.....

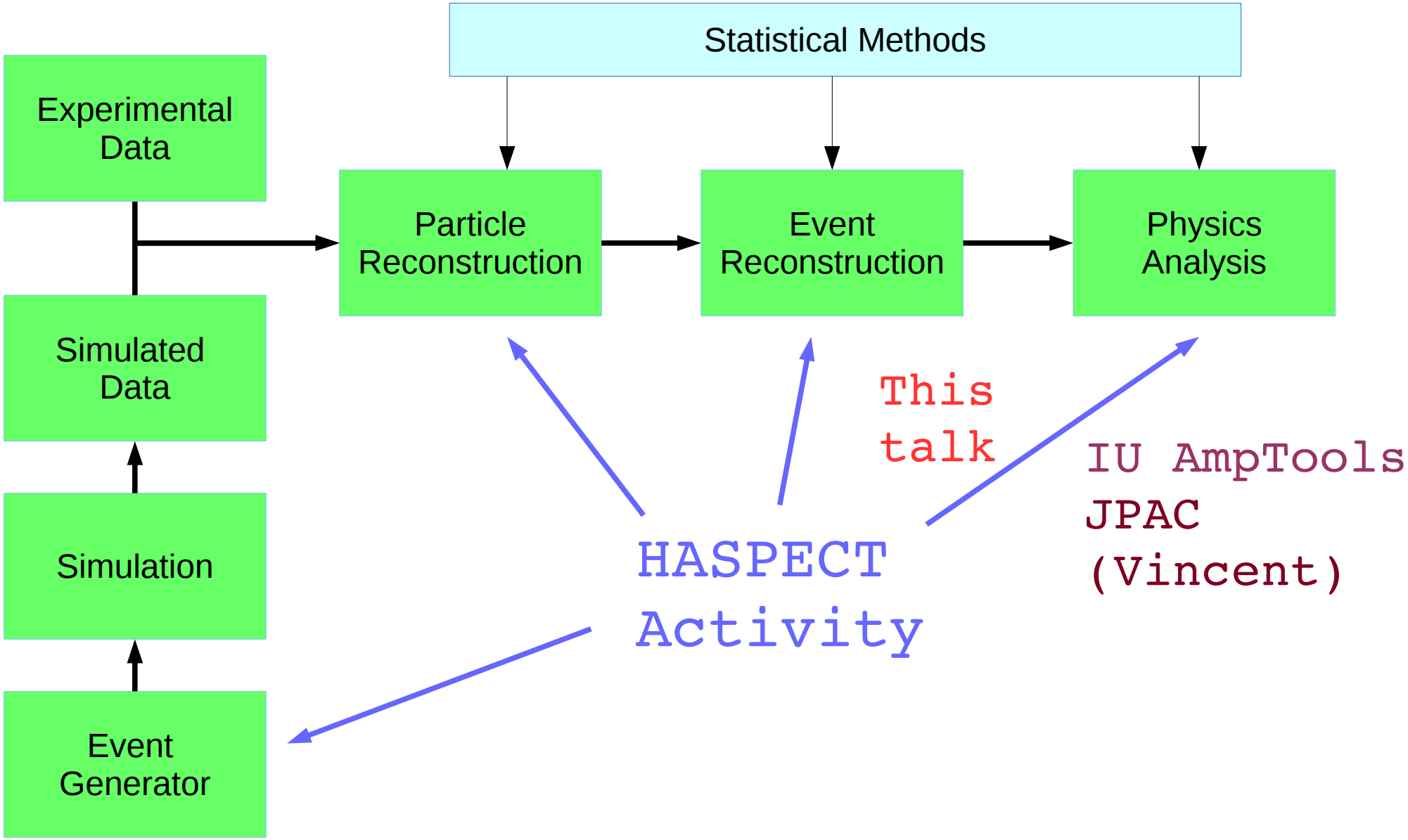
Theoretical support:
A.Szczepaniak (IU/JPAC), V.Mathieu (IU), E.Santopinto (INFN-GE), A.Vassallo (GE), J.Ferretti (UMAS)

Experimental Analysis:
M.Battagleiri, R.deVita, A.Celentano, S.Fegan (INFN-GE), A. Filippi (INFN-TO), D.Glazier(Glasgow), S.Hughes (Edinburgh), K.Hicks (OhioU), S.Lombardo (Cornell), A.Rizzo (RomaTV), I Stankovich (Edinburgh), L.Zana (Edinburgh)

Use of Software in hadron physics



Use of Software in hadron physics



HASPECT Event Reconstruction

Provide code to handle routine tasks
allowing procedures to become standardised

- Input/Output/Interfacing
- Histogramming
- Particle/reaction identification
- Event weighting

Maintain normal ROOT flexability for users

Users shift to physics and systematic studies

Promote full potential of ROOT

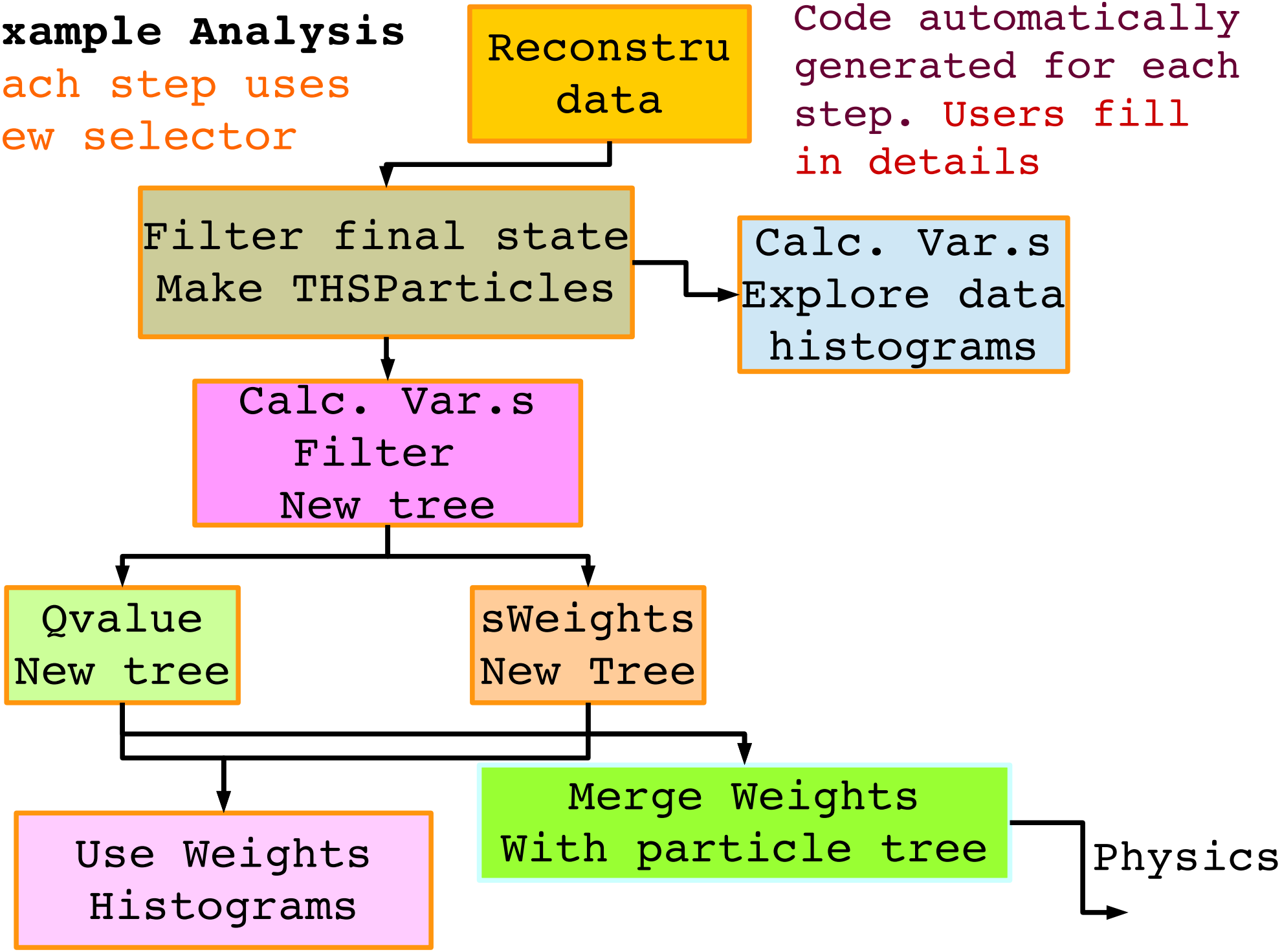
- Based on TSelector Tree analysis class
- Use of TEntryList class to prevent duplicating data
- ROOT system takes care of compilation and configuration
- Parallel ROOT Facility (PROOF)
- Statistical Analysis Packages (RooFit/Stats)

https://github.com/HASPECT/Events

Example Analysis

Each step uses new selector

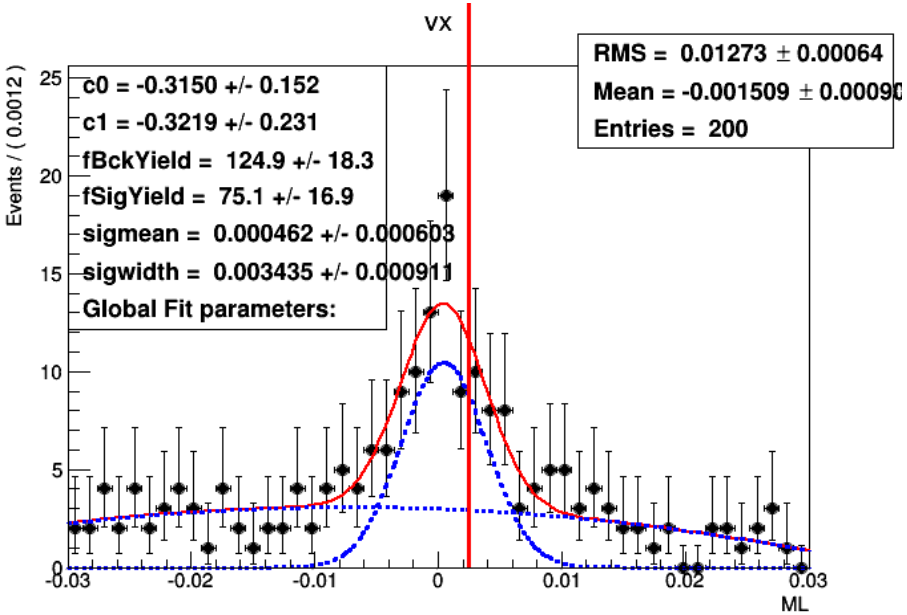
Code automatically generated for each step. Users fill in details



Event Weighting : Qvalue

Developed Mike Williams (CMU) for CLAS analysis
Look for N nearest events in kinematic space
Fit discriminatory variable for signal and back.
e.g. missing mass
Qval = S/(S+B)
Qval can then be used to weight events

HASPECT
*Selector class inherits additional Qvalue class
*Use RooFit event-by-event maximum likelihood
*Near. Neigh. saved
*Limit NN search with TEntryList



Event Weighting : sWeights

M. Pivk,F.R. Le Diberder,Nucl.Inst.Meth.A 555, 356-369, 2005

Given discriminatory PDF for signal and background calculates weight :

$$ {}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_S} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_S} N_k f_k(y_e)} $$

N_s = Number of species

f_k = PDF for species k

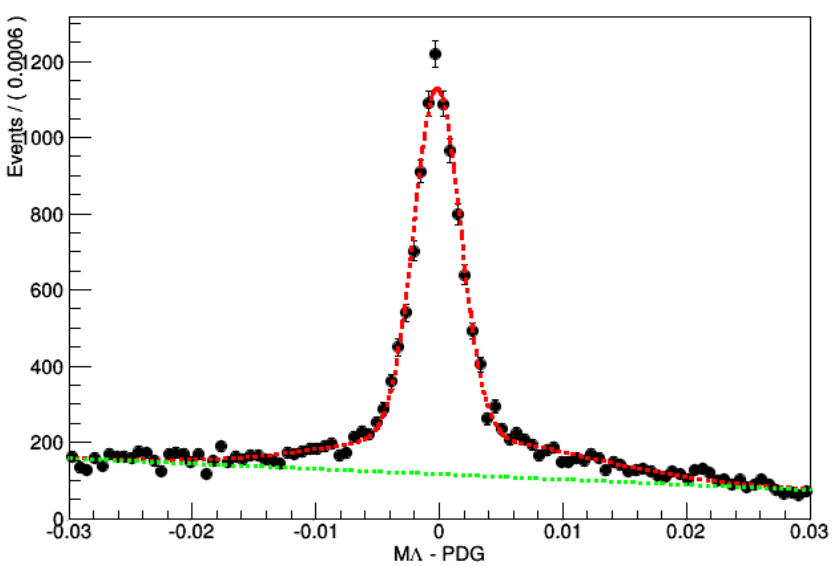
N_k = Yield for species k

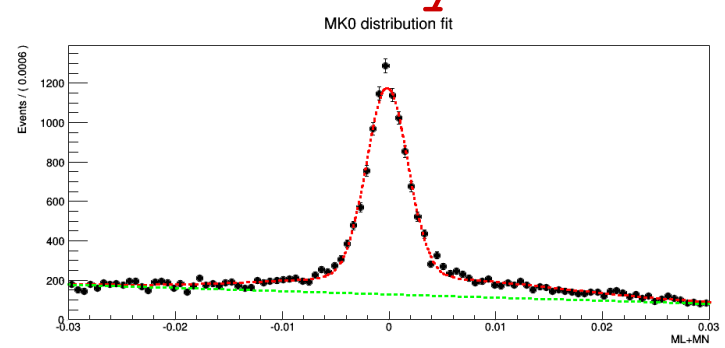
V = covariance matrix

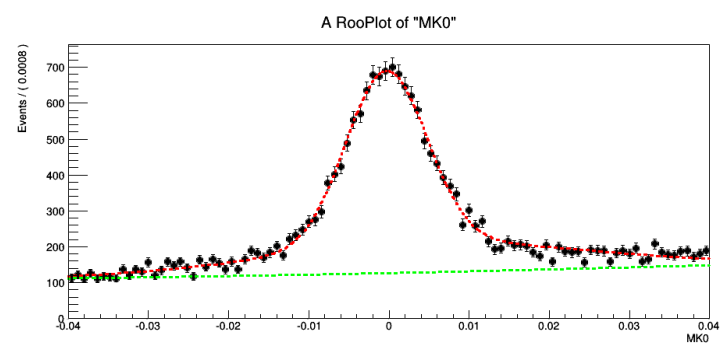
Part of RooStats(used here)

Can include multiple signal and background species

Can fit multidimensional discriminatory PDF







Test on "perfect" data (simulation)

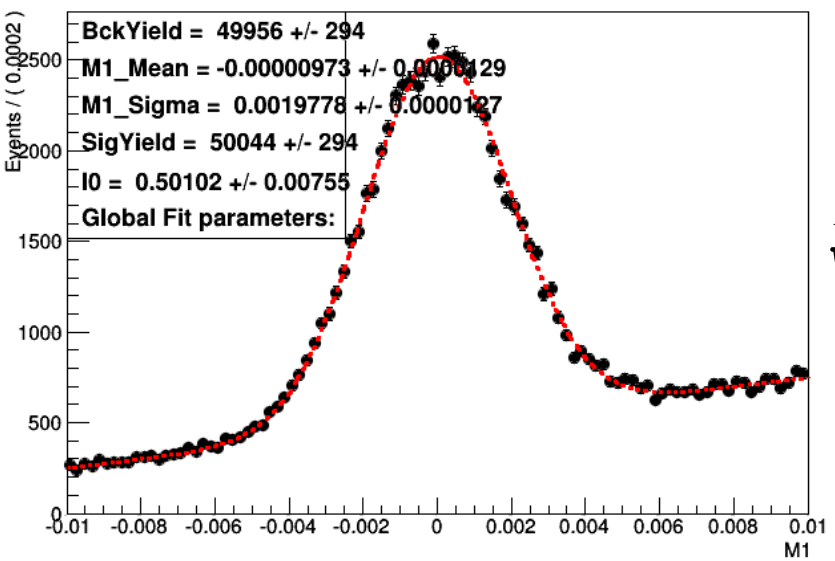
Discriminatory variable M1

Signal Gaus(0,0.002) Background Linear

Weighted variable M2

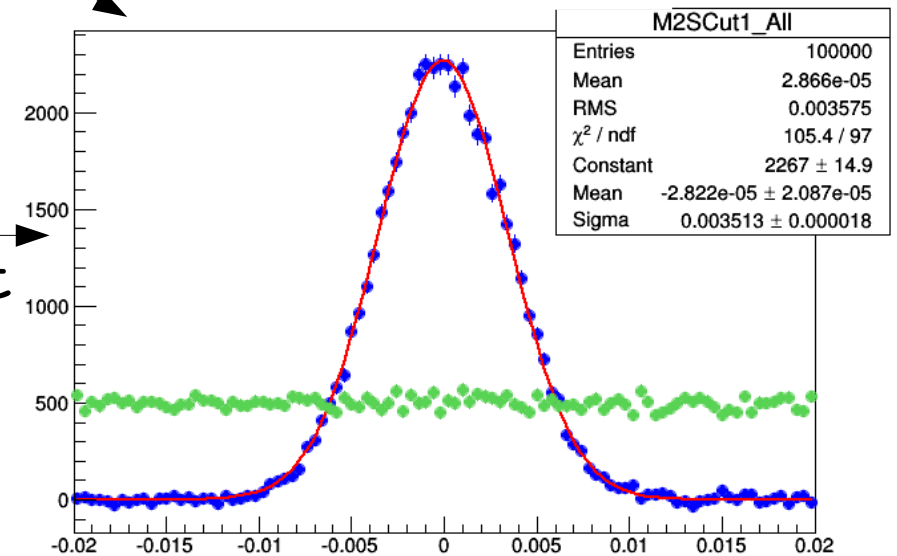
Signal Gaus(0,0.0035) Background Flat

M1



Calc. Weight

M2



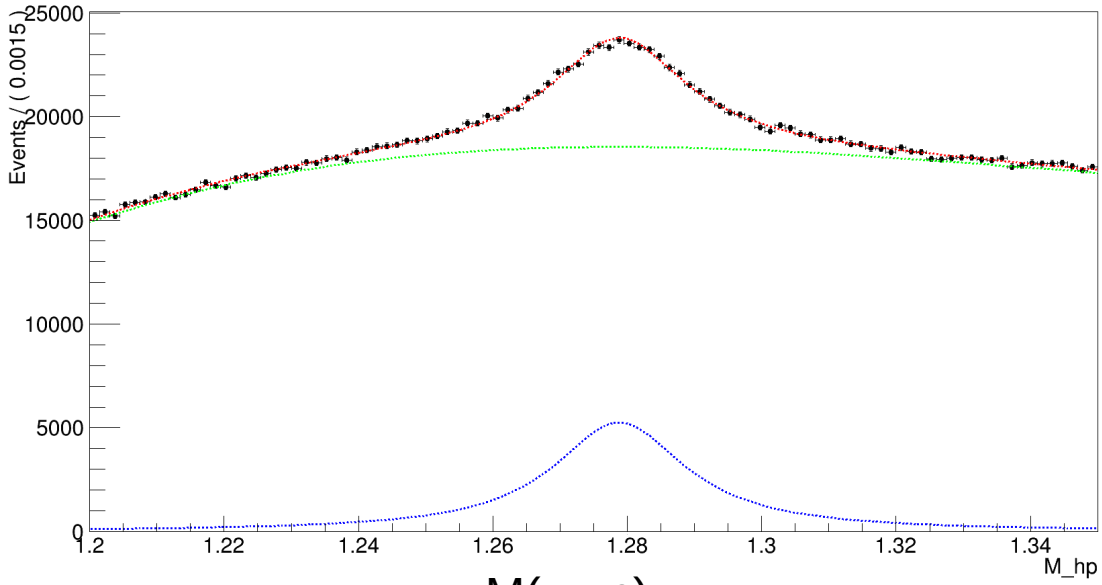
Discriminatory Variable

Result

sWeights correctly reproduces M2 signal and background shape and uncertainties

Example CLAS Analysis

$\gamma p \rightarrow f1(1285)\ p$
$\rightarrow \pi^+\pi^-(\eta)p$



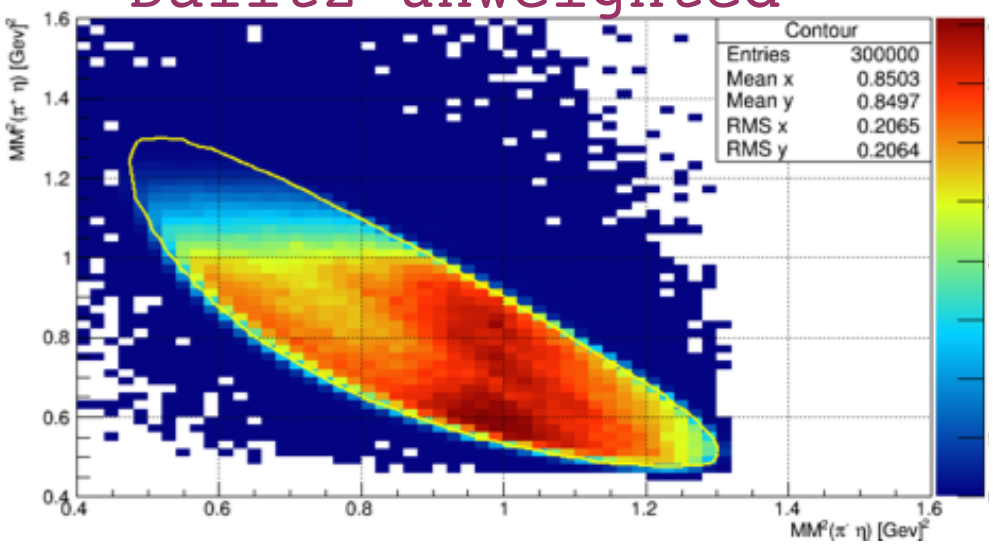
M(ππη)

Goal: Fit Veneziano(B4) amplitude to 3 meson decay

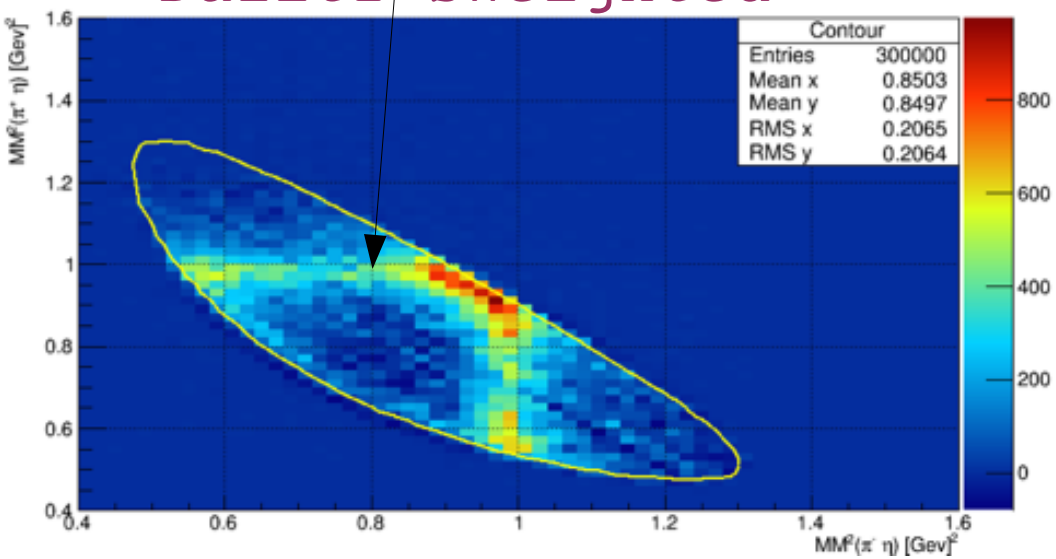
(Alessandro Rizzo)

Clear $f1 \rightarrow \pi a_0(980)$

Dalitz unweighted

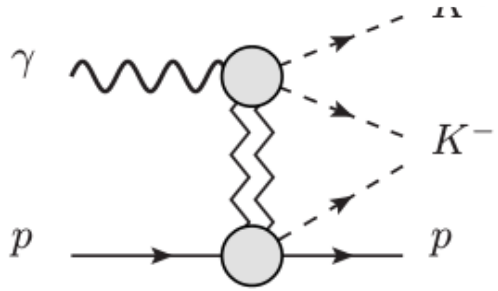


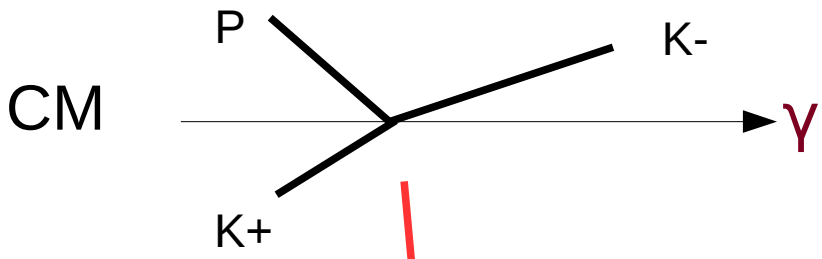
Dalitz sWeighted

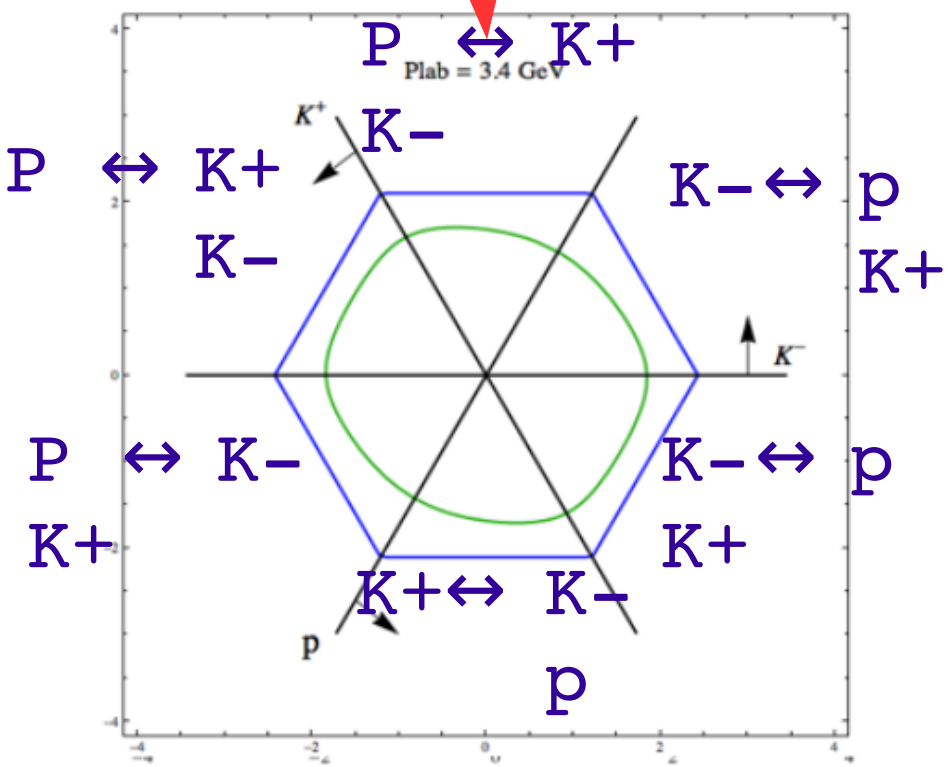


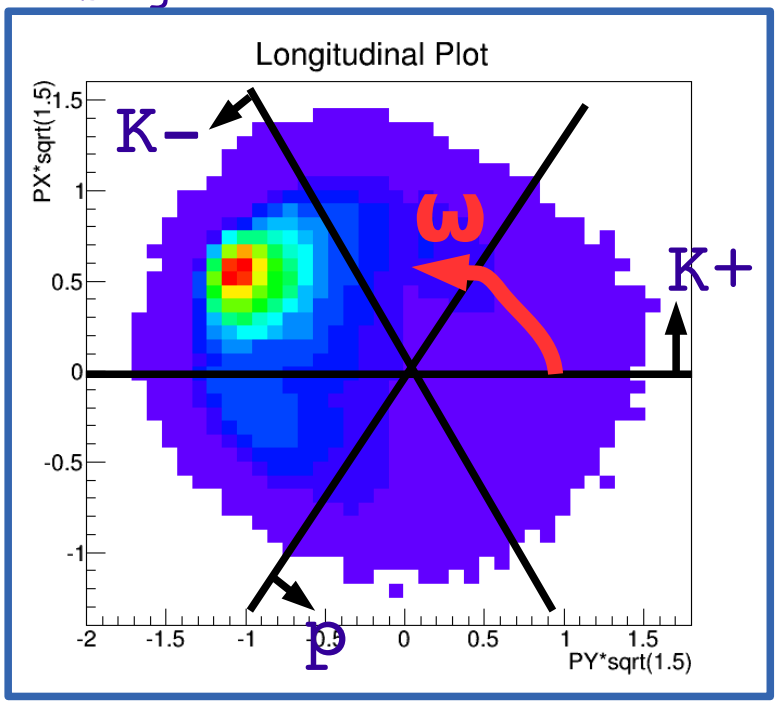
Van Hove Plots (Longitudinal)

Example 3-3.8GeV γp→K+K-p
CLAS g11 dataset







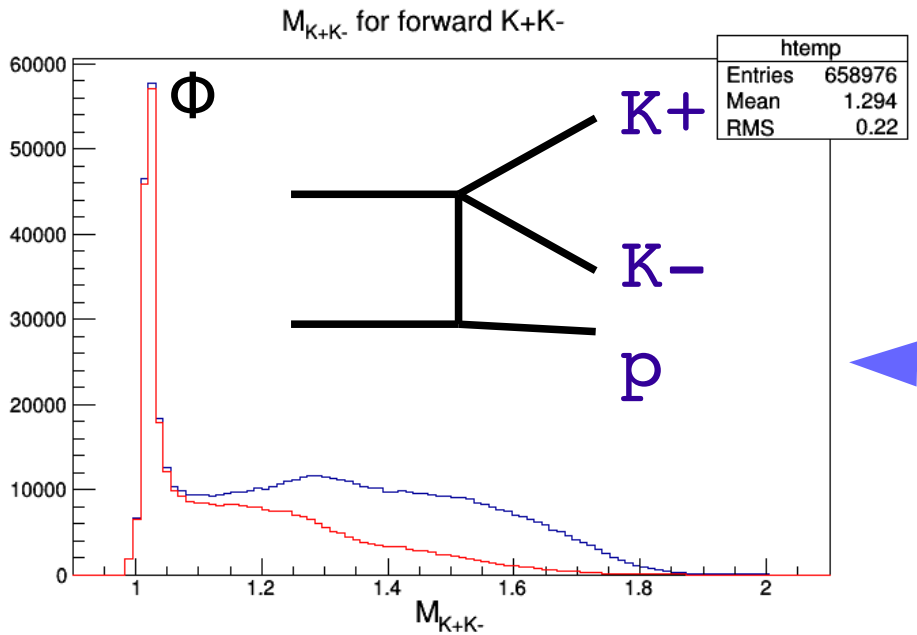


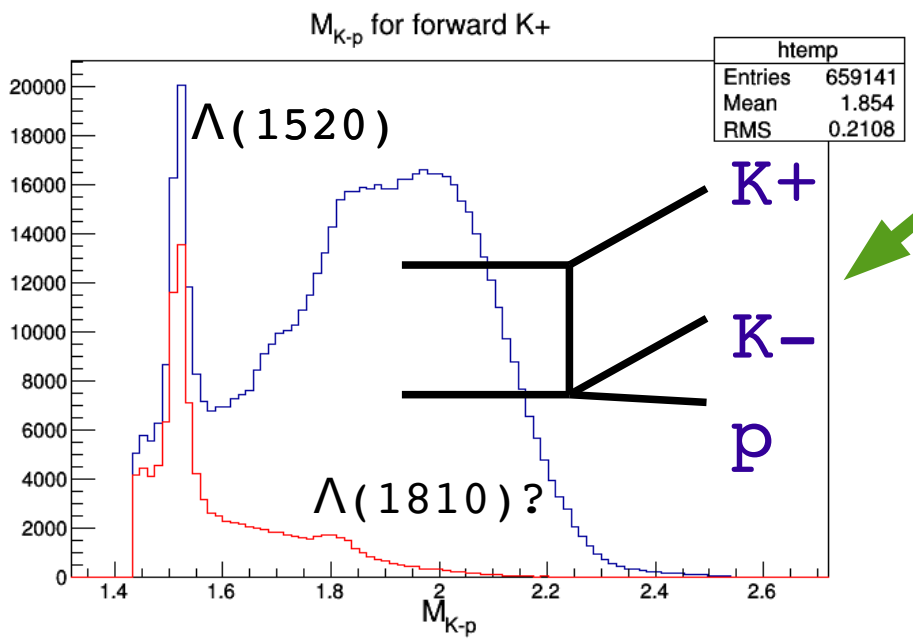
$$p_{K^+L} = \sqrt{\frac{2}{3}} q \sin\omega,$$

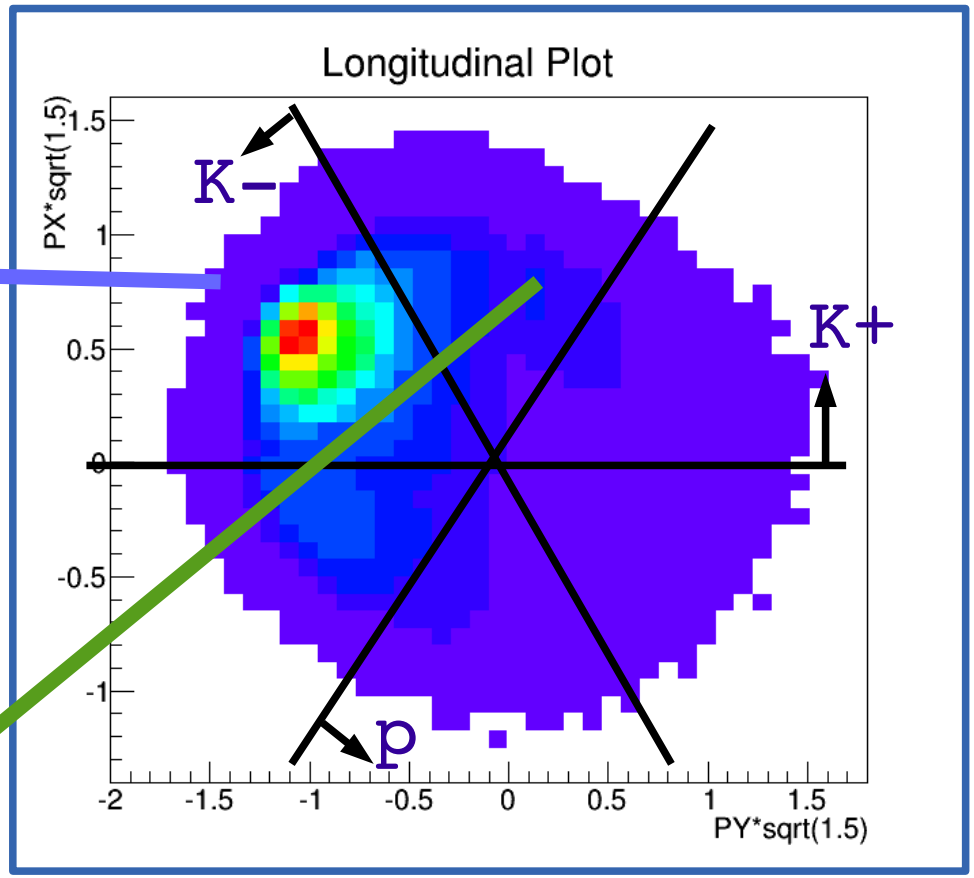
$$p_{K^-L} = \sqrt{\frac{2}{3}} q \sin\left(\frac{2}{3}\pi + \omega\right),$$

$$p_{PL} = \sqrt{\frac{2}{3}} q \sin\left(\frac{4}{3}\pi + \omega\right).$$

Example γp→K+K-p at around 3-3.8GeV

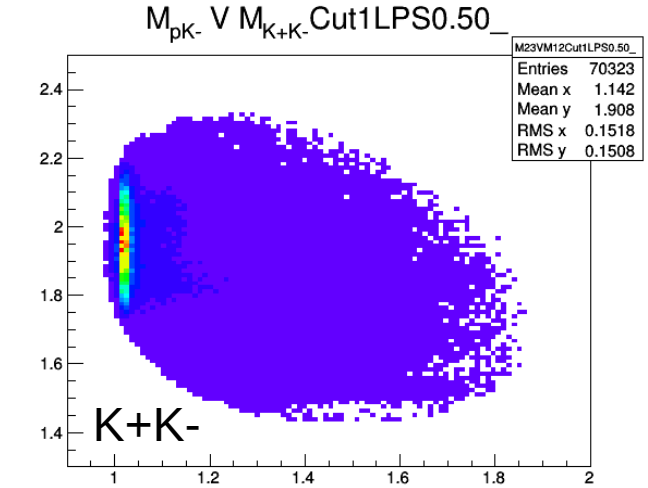


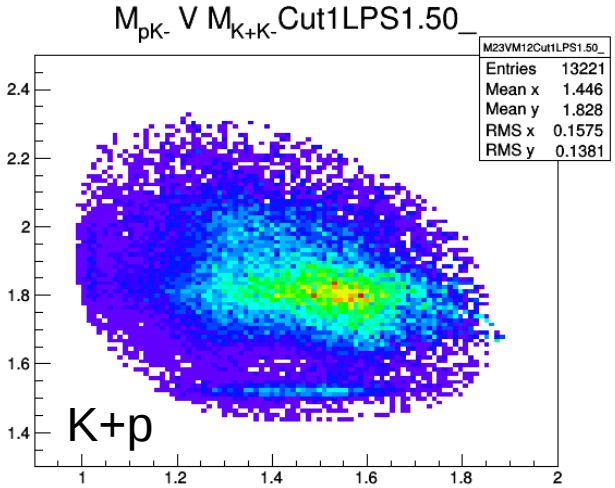


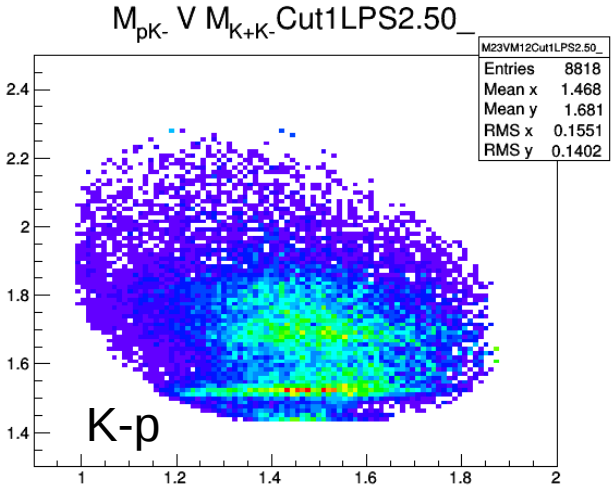


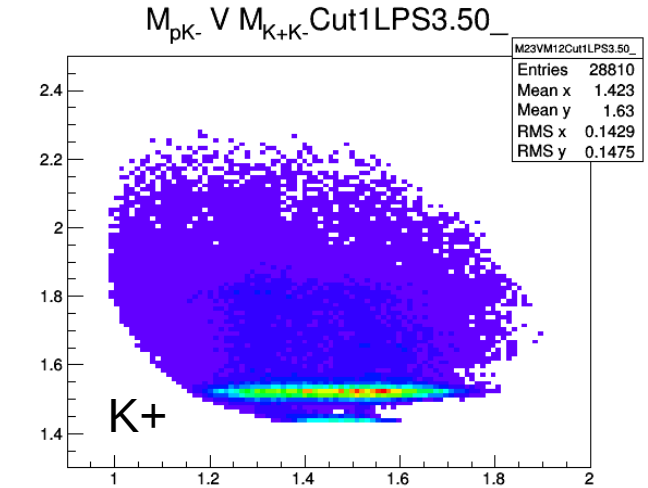
All Events

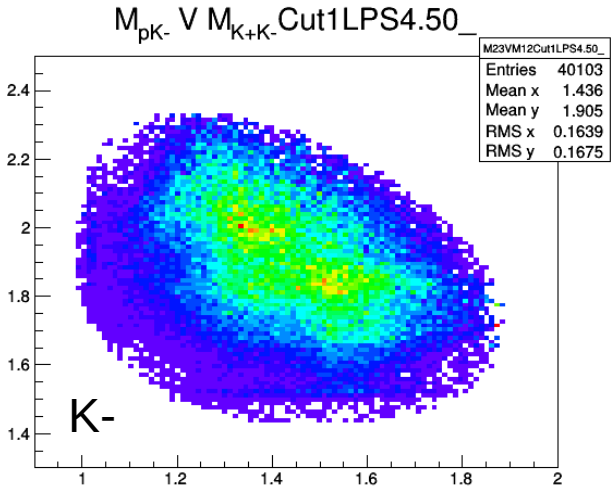
Cut on Longitudinal Plot sector

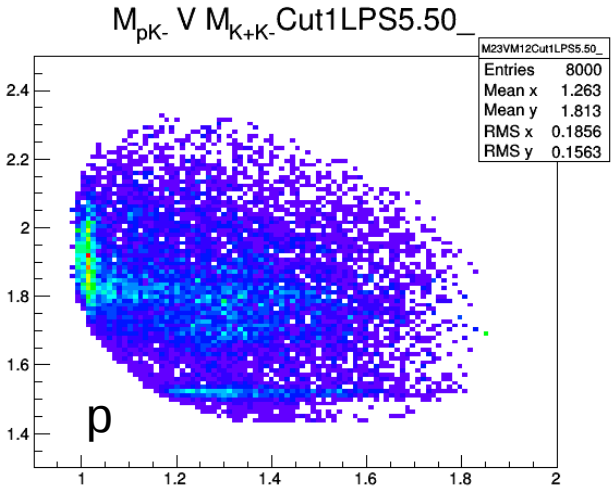

M_{pK-} V M_{K+K-}Cut1LPS0.50_
M23VM12Cut1LPS0.50_
Entries 70323
Mean x 1.142
Mean y 1.908
RMS x 0.1518
RMS y 0.1508
K+K-

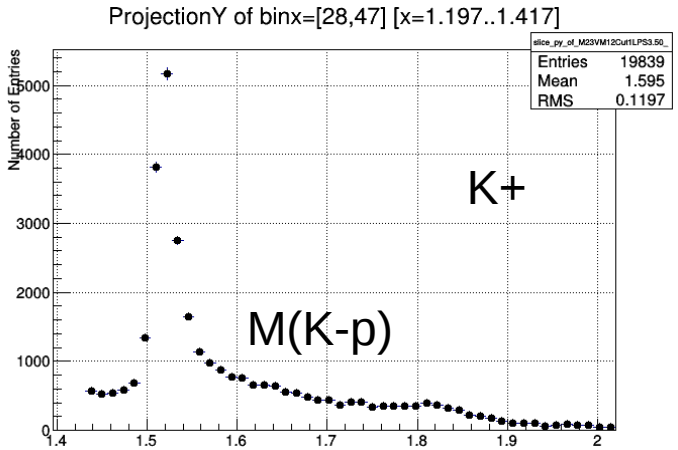

M_{pK-} V M_{K+K-}Cut1LPS1.50_
M23VM12Cut1LPS1.50_
Entries 13221
Mean x 1.446
Mean y 1.828
RMS x 0.1575
RMS y 0.1381
K+p

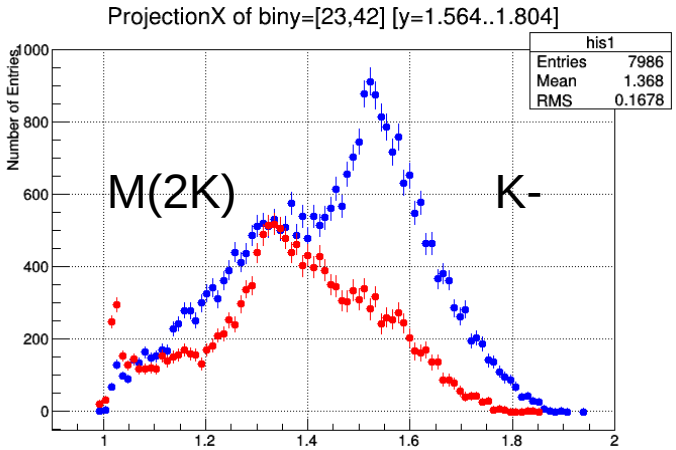

M_{pK-} V M_{K+K-}Cut1LPS2.50_
M23VM12Cut1LPS2.50_
Entries 8818
Mean x 1.468
Mean y 1.681
RMS x 0.1551
RMS y 0.1402
K-p

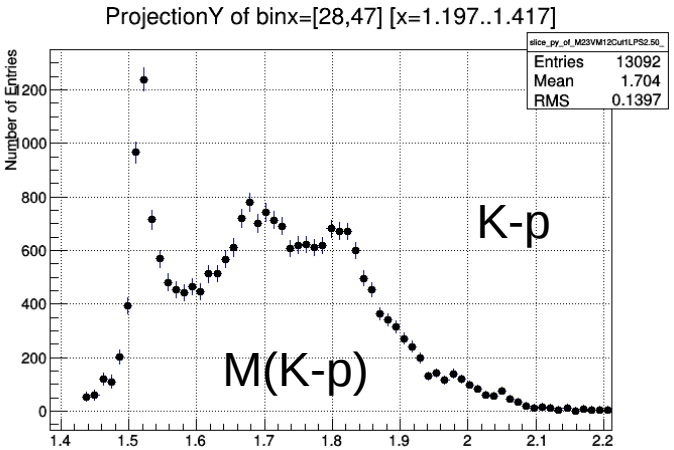

M_{pK-} V M_{K+K-}Cut1LPS3.50_
M23VM12Cut1LPS3.50_
Entries 28810
Mean x 1.423
Mean y 1.63
RMS x 0.1429
RMS y 0.1475
K+


M_{pK-} V M_{K+K-}Cut1LPS4.50_
M23VM12Cut1LPS4.50_
Entries 40103
Mean x 1.436
Mean y 1.905
RMS x 0.1639
RMS y 0.1675
K-

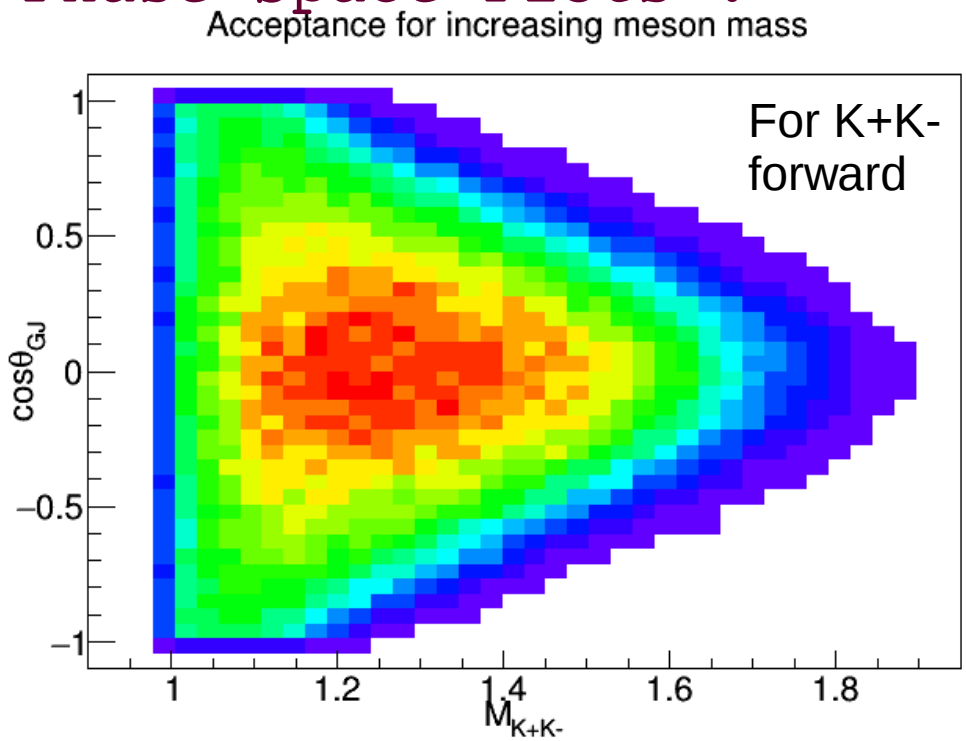

M_{pK-} V M_{K+K-}Cut1LPS5.50_
M23VM12Cut1LPS5.50_
Entries 8000
Mean x 1.263
Mean y 1.813
RMS x 0.1856
RMS y 0.1563
p

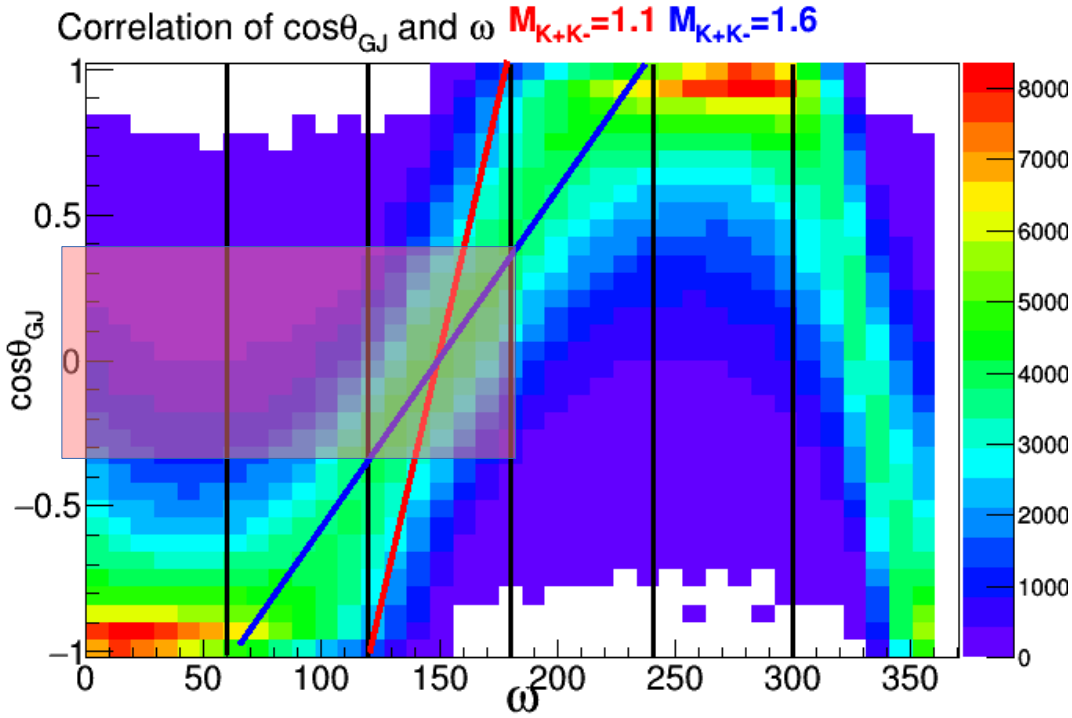

ProjectionY of binx=[28,47] [x=1.197..1.417]
slice_py_of_M23VM12Cut1LPS3.50_
Entries 19839
Mean 1.595
RMS 0.1197
Number of Entries
K+
M(K-p)


ProjectionX of biny=[23,42] [y=1.564..1.804]
his1
Entries 7986
Mean 1.368
RMS 0.1678
Number of Entries
M(2K)
K-


ProjectionY of binx=[28,47] [x=1.197..1.417]
slice_py_of_M23VM12Cut1LPS2.50_
Entries 13092
Mean 1.704
RMS 0.1397
Number of Entries
K-p
M(K-p)

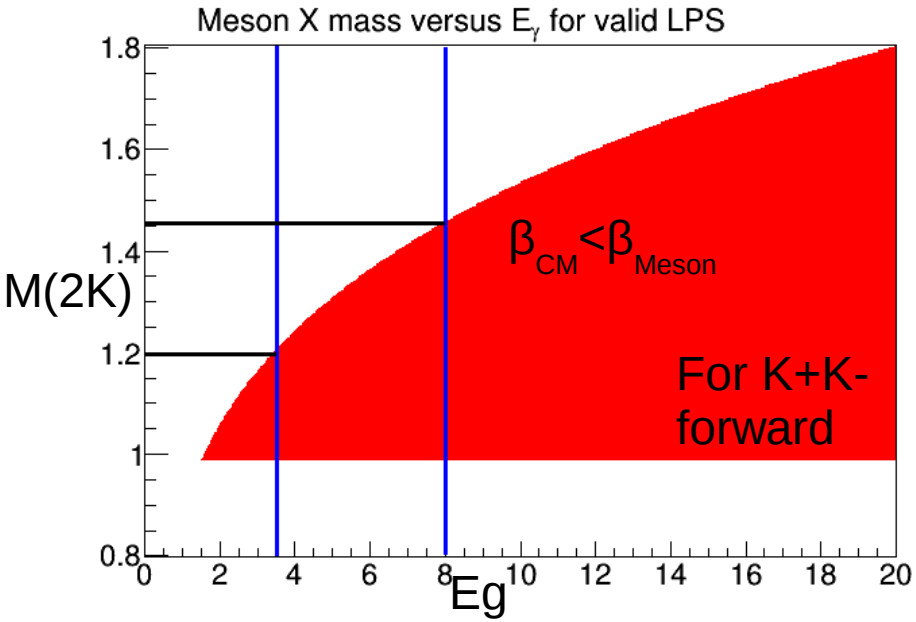
Larger Mass 2K mesons will have lower CM momenta
Decay products can decay back into different sector
Phase Space Plots :





K- back
From decay

K+ back
from decay
of meson



M(2K) = 1-1.2 OK, but...
M(2K) > 1.2 has limited θ_{GJ}

Monte-Carlo Sampling of likelihood

Minuit Maximum Likelihood
- Single solution

We often have local maxima
How to choose initial parameters?
How to judge goodness of fit?
Implement Occam's Razor?

MCMC
- Samples full likelihood
- Not very efficient sampling
- Only finds unimodal solution
- Difficult to calculate evidence

Nested Sampling for General Bayesian Computation –
- J. Skilling, 2006, *International Society for Bayesian Analysis*
  - More efficient sampling
  - Intrinsically calculates evidence

model selection via Bayes factor
+Occams Razer

MultiNest
- F.Feroz and M.P. Hobson, 2007,*Cambridge, arXiv:0704.3704v3*
- Finds many maxima and the evidence for each

Bayesian Statistics is used for parameter estimation and hypothesis testing.

INPUT OUTPUT

Likelihood x *Prior* = *Evidence* x *Posterior*

$$P(\mathbf{D}|\boldsymbol{\theta},H) \times P(\boldsymbol{\theta}|H) = P(\mathbf{D}|H) \times P(\boldsymbol{\theta}|\mathbf{D},H)$$

$$L(\theta) \times \pi(\theta)d\theta = Z \times p(\theta)d\theta$$

Where **D** is the data set, **θ** is a parameter vector, H is a model and

$$Evidence = Z = \int L(\theta)\pi(\theta)\,d\theta$$

Likelihood integrated over the prior distribut.

Nested Sampling estimates evidence and finds likelihood maxima.

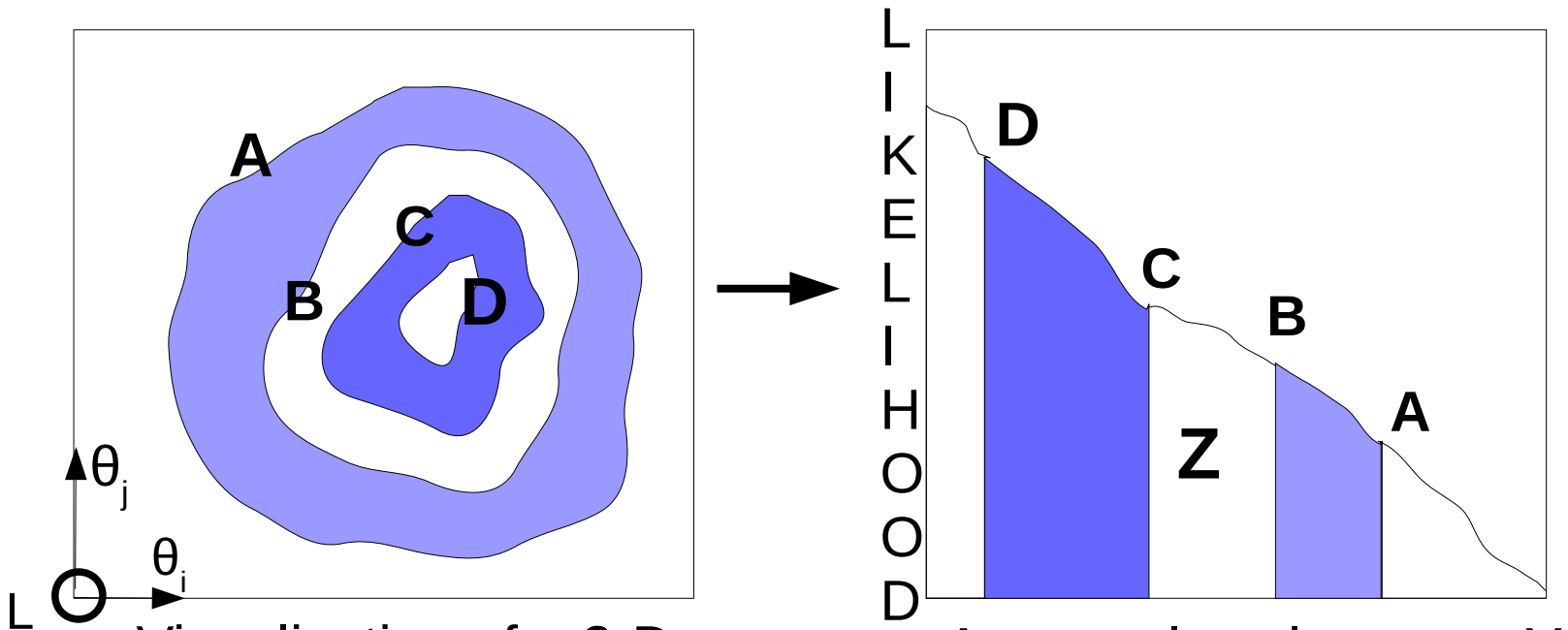
Define accumulated mass

$$X = \int_{L(\theta)>\lambda} \pi(\theta)\, d\theta$$

Evidence can be evaluated as a 1D integral of likelihood over the prior accumulated mass

$$Z = \int L(X)\, dX$$

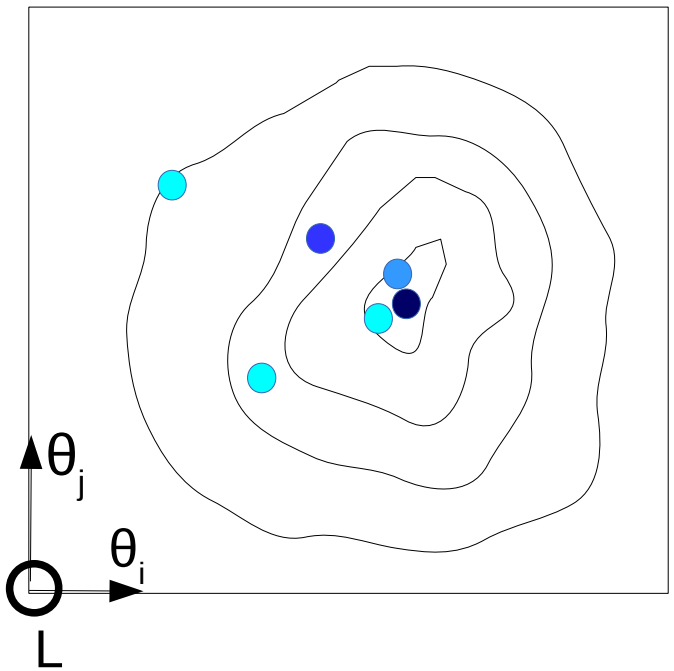
N “live points” maintained and one with lowest likelihood is replaced

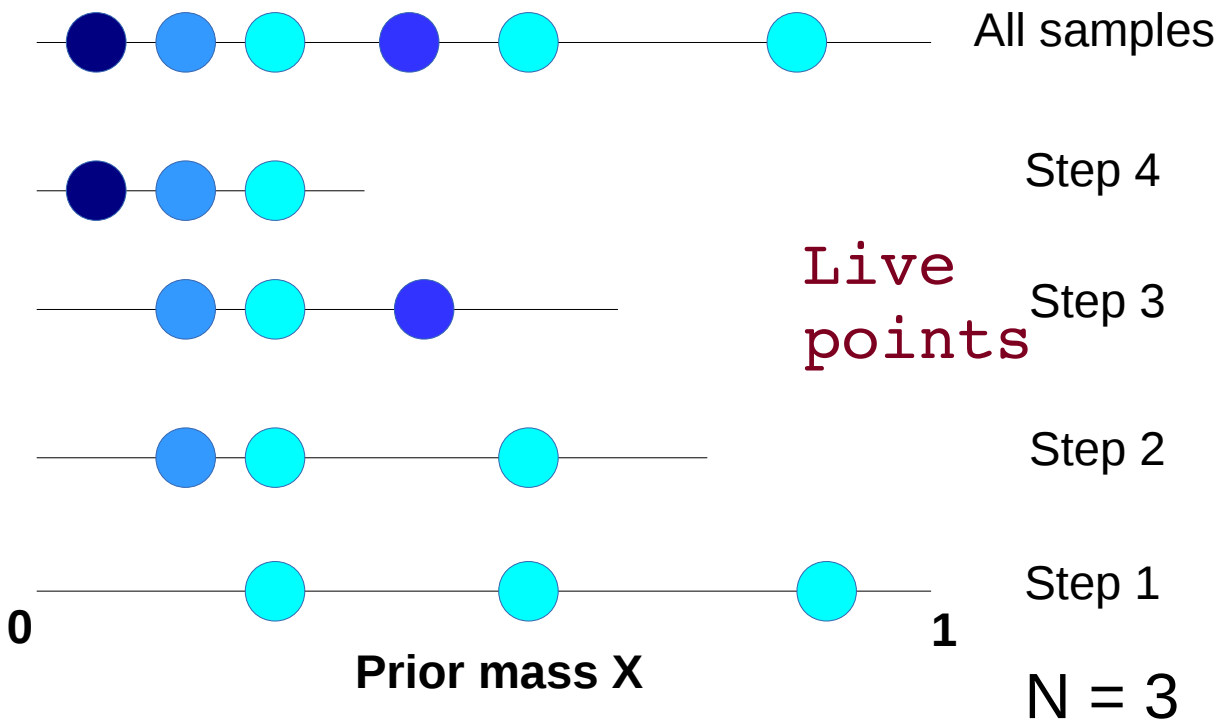


Visualization of a 2-D parameter space

Accumulated mass - X

Nested Sampling procedure involves points exploring the likelihood.





Note that only one new point needs to be calculated at each iteration, the N points at iteration *i* are the active live points.

Rejected points are kept to give the posterior distribution (All samples)

MultiNest in IUAmpTools

R. Mitchell, IndianaU
M. Shepherd
H. Matevosyan
L. Gibbons

Calculate Intensity in terms of production and decay amplitudes

Constructed by user

kinematics derived from 4-vectors

decay amplitudes (from theory)

$$I(\Omega) = \sum_{\alpha} \left| \sum_{\beta} V_{\alpha\beta} A_{\alpha\beta}(\Omega) \right|^2$$

incoherent sum

coherent sum

production amplitudes (complex fit parameters)

Minimise :

From data

From Simulation

$$-2 \ln L = -2 \sum_{i=1}^{N_{\text{observed}}} \ln(I'(\Omega_i)) + \frac{2}{N^{\text{MC}}_{\text{generated}}} \sum_{i=1}^{N^{\text{MC}}_{\text{accepted}}} I'(\Omega_i)$$

AmpTools Dalitz Tutorial

$X(3000) \rightarrow P_1(200)P_2(200)P_3(200)$

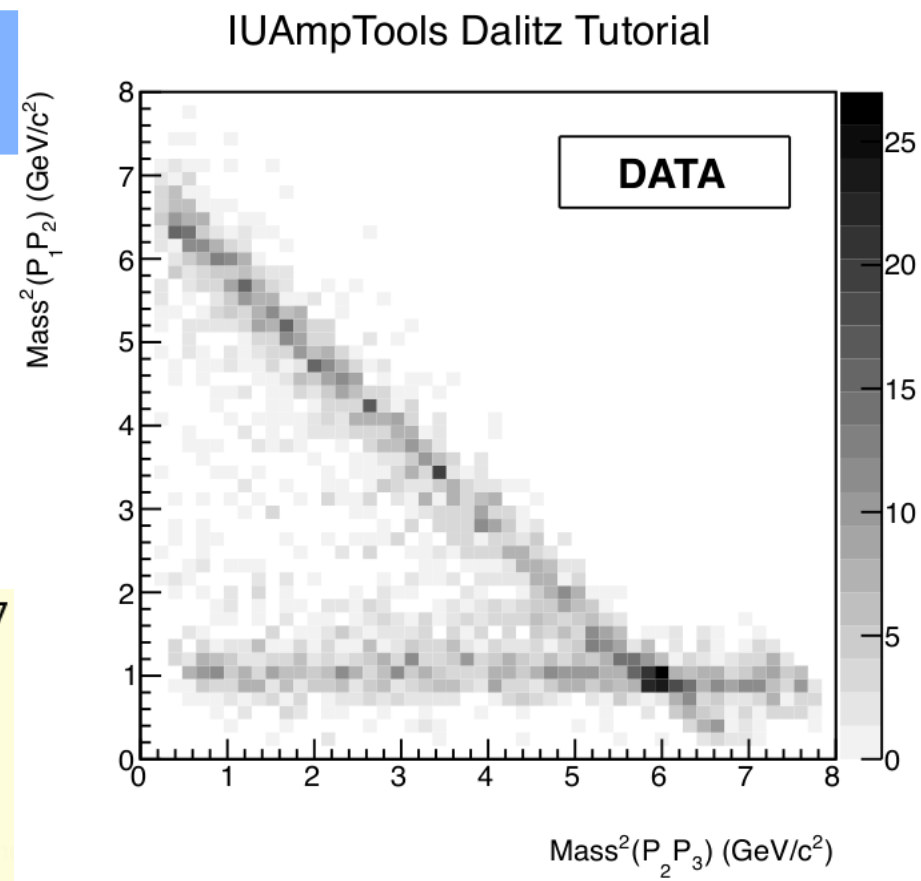
Amplitude $A_{\alpha\beta}$: Breit-Wigner Isobars in P_1P_2 and P_1P_3

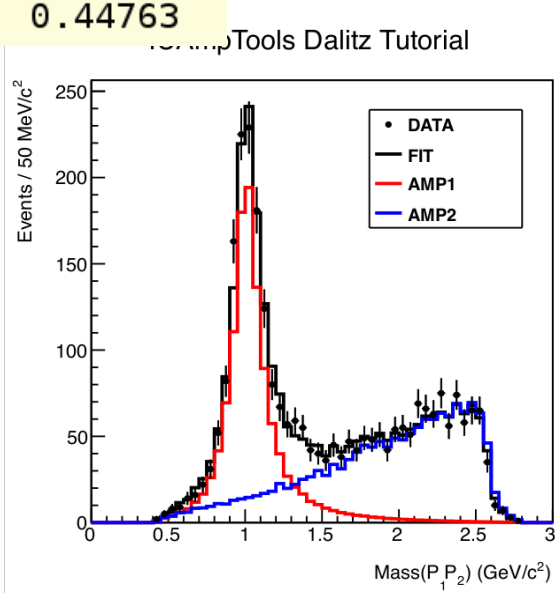
```
FCN=-46559.8 FROM MIGRAD     STATUS=CONVERGED     197

                    EDM=6.19563e-06    STRATEGY= 1
 EXT PARAMETER
 NO.   NAME                VALUE            ERROR
   1 dalitz::s1::R12_re         30.98           0.4832
   2 dalitz::s1::R12_im       -3.0114           1.8561
   3 dalitz::s1::R13_re        30.639          0.44763
```

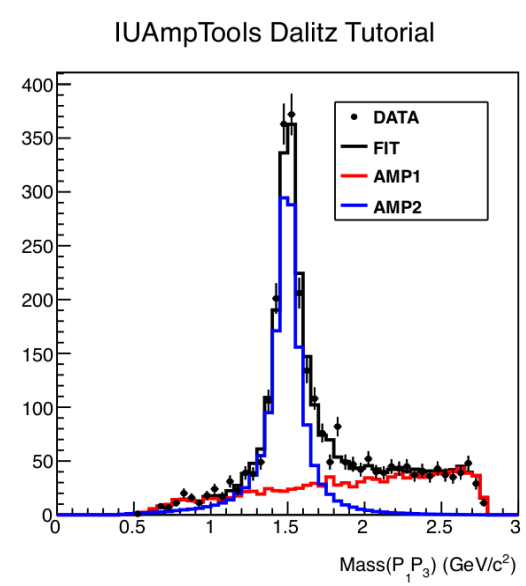
Fit "prod, Amps" $V_{\alpha\beta}$

Minuit result in 0.03s

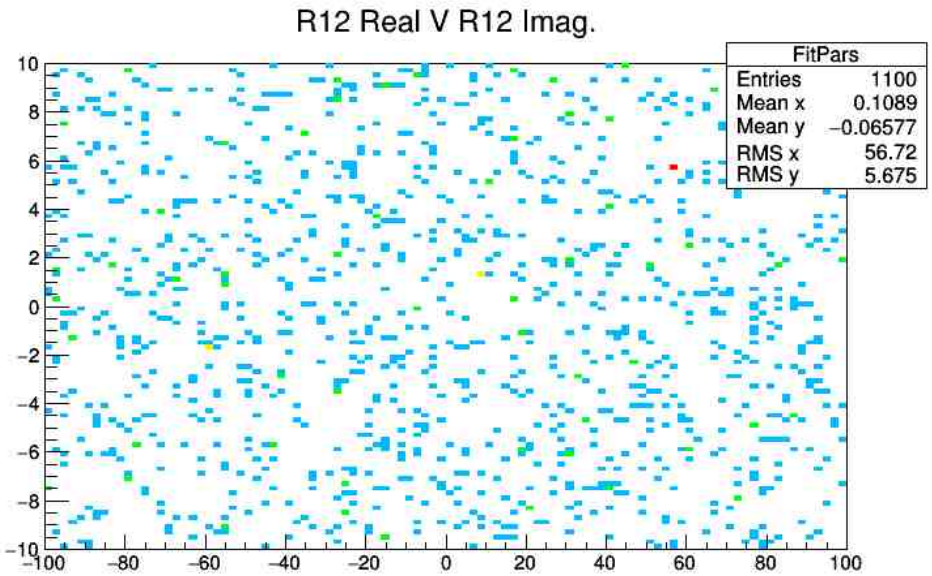
2800 generated data events

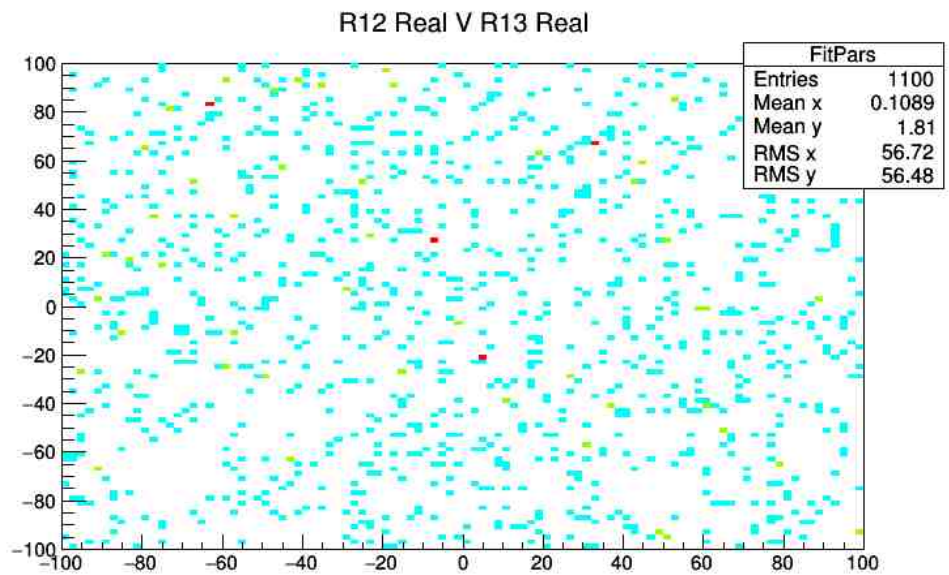


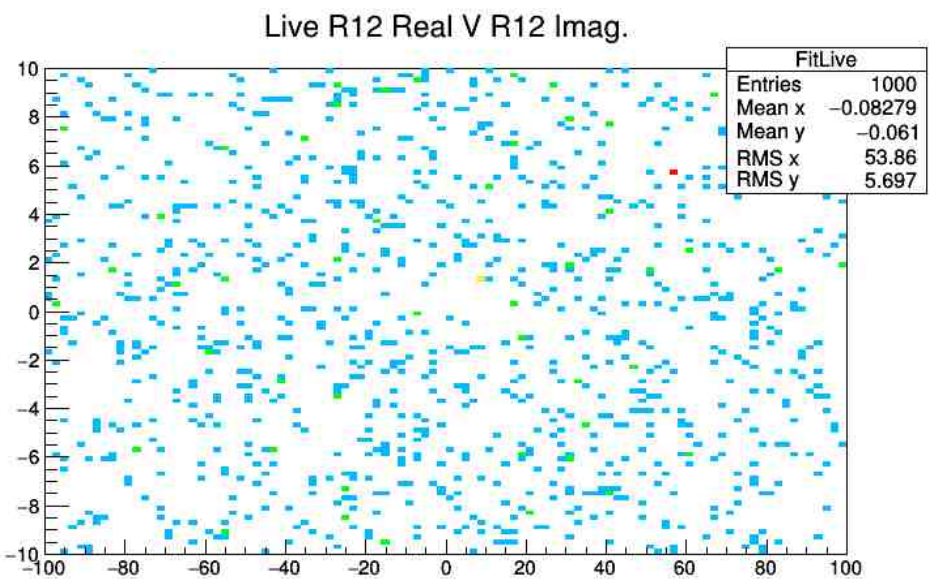


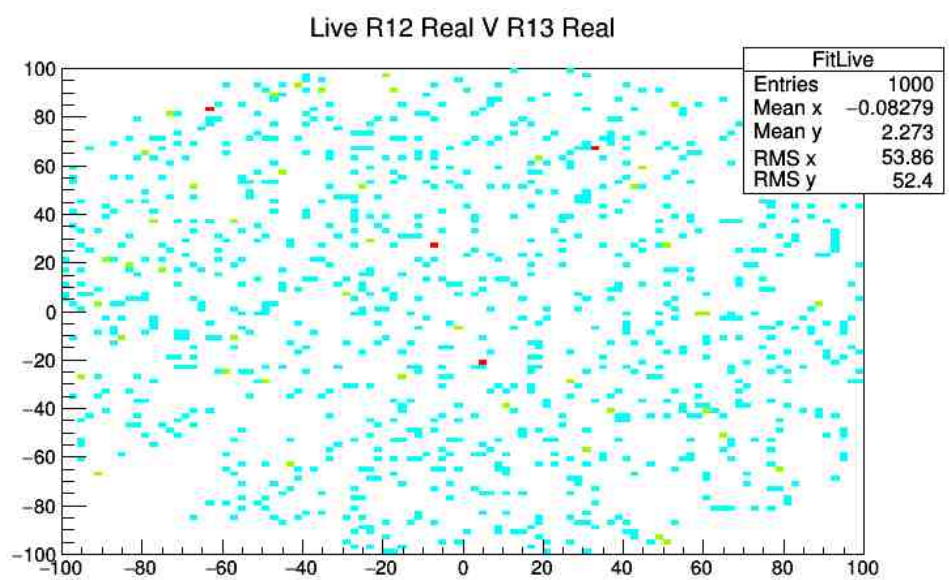


MultiNest Dalitz Fit, Posterior and Live Points



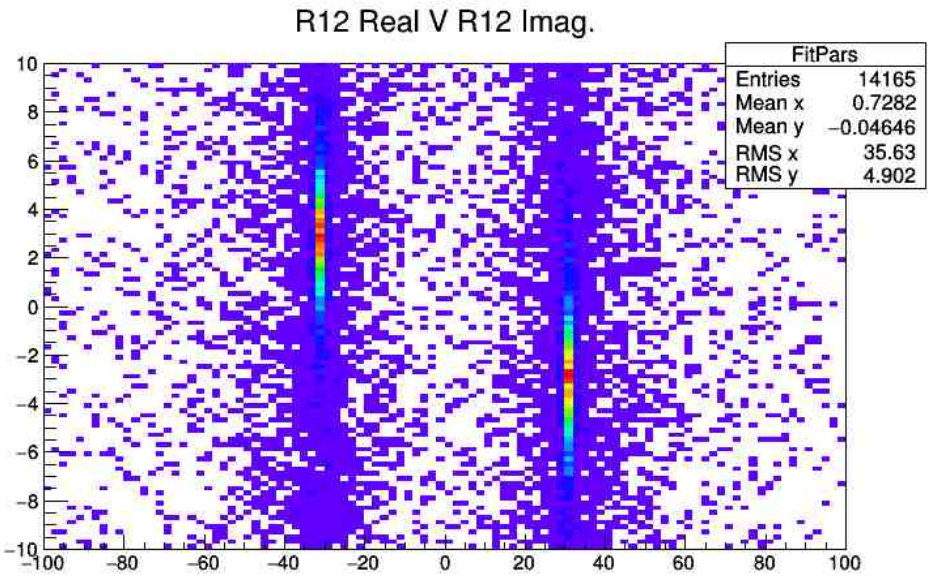


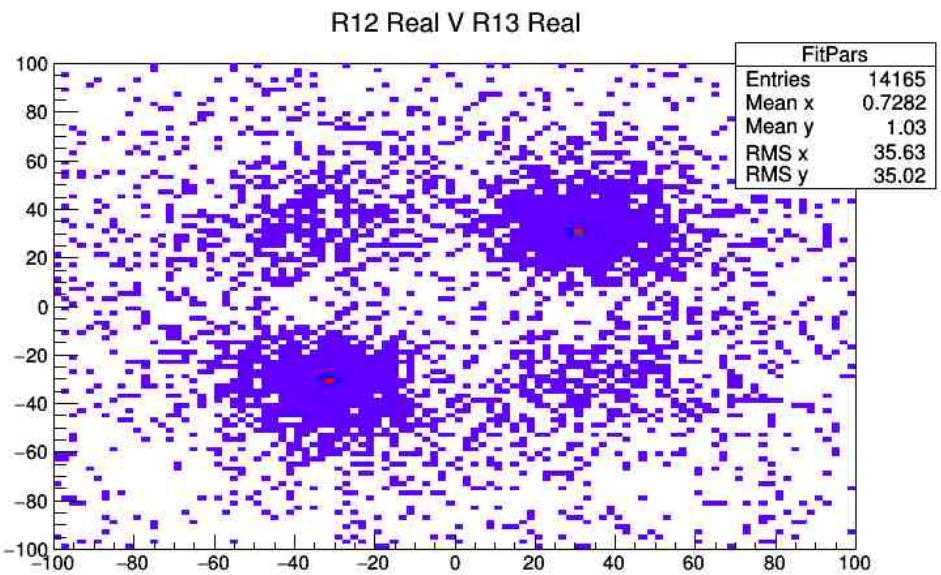


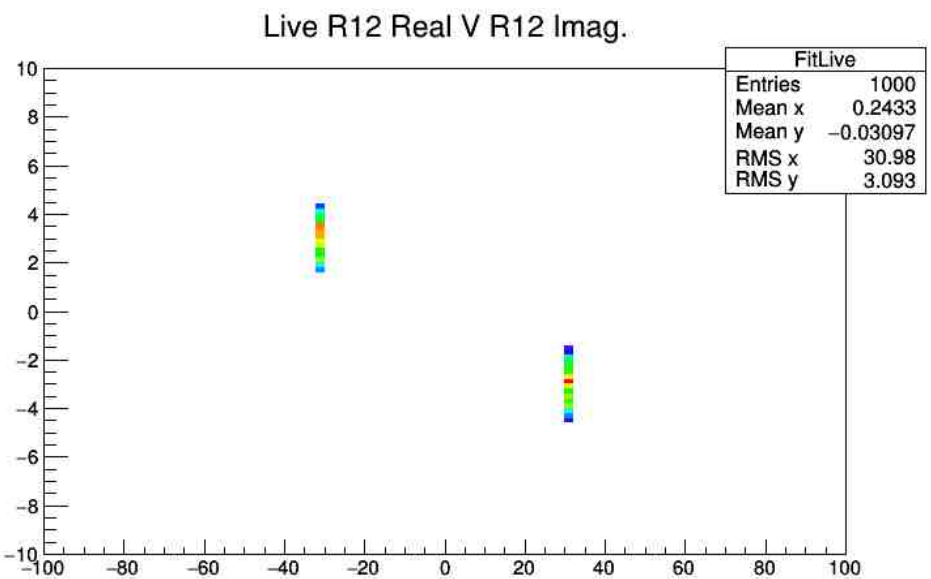


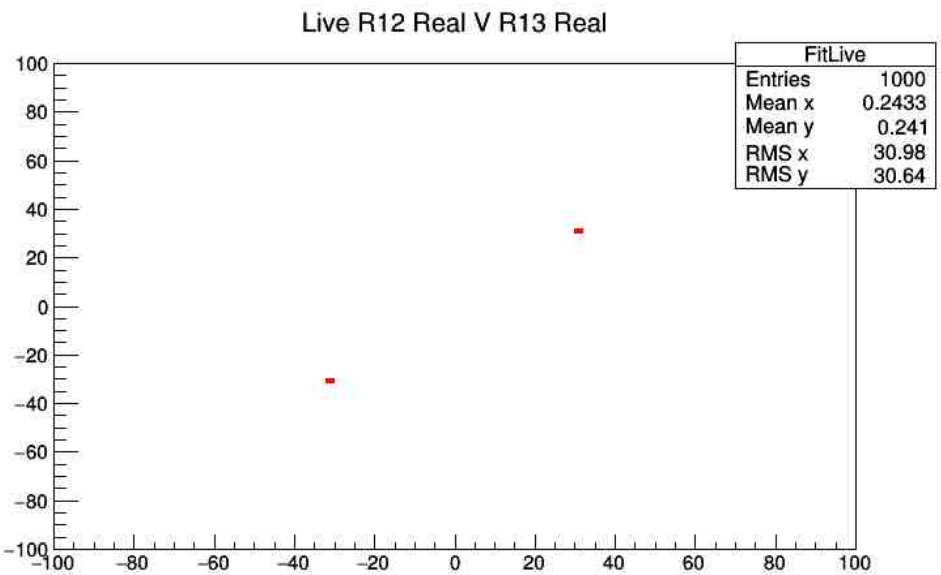
http://www.nuclear.gla.ac.uk/~dglazier/multinest/Dalitz1.mov

MultiNest Dalitz Fit, Posterior and Live Points

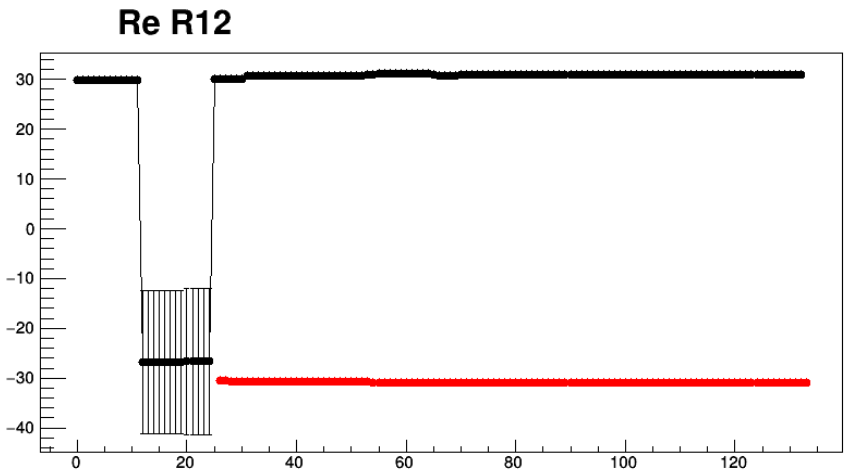


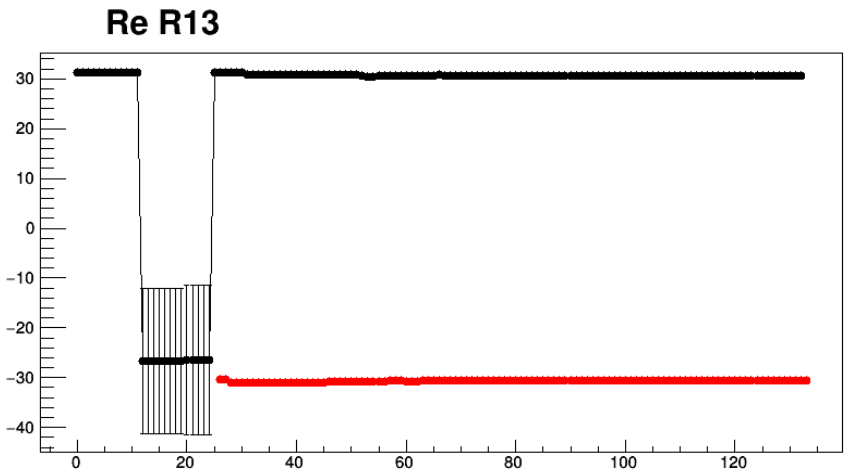


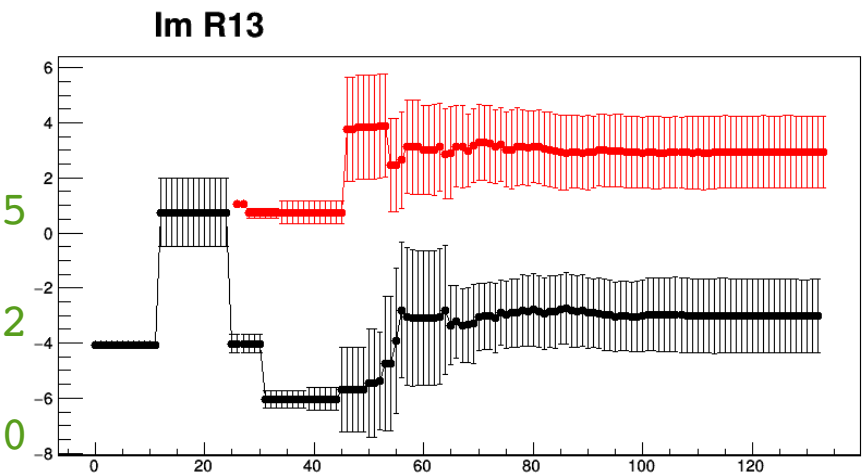




MultiNest Dalitz Fit
– 2 solutions, Evolution with iteration number







Mode 1	Mode 2
Re R12 =30.96∓0.34	Re R12=-30.96∓0.35
Re R13 =30.64∓0.32	Re R13=-30.64∓0.32
Im R13 = 3.03∓1.33	Im R13 = 2.94∓1.30

CPU time for 100 Live Points = 0.7s

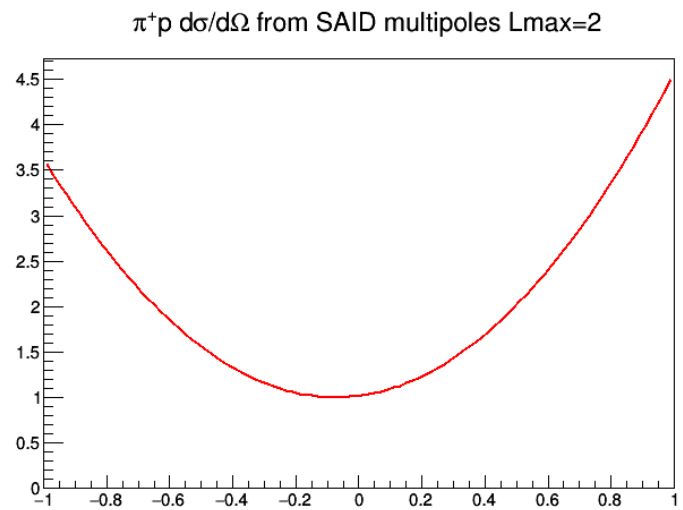
Start with a simple case $\pi\vec{N}\rightarrow\pi N$

Use SAID PW
At 200 MeV

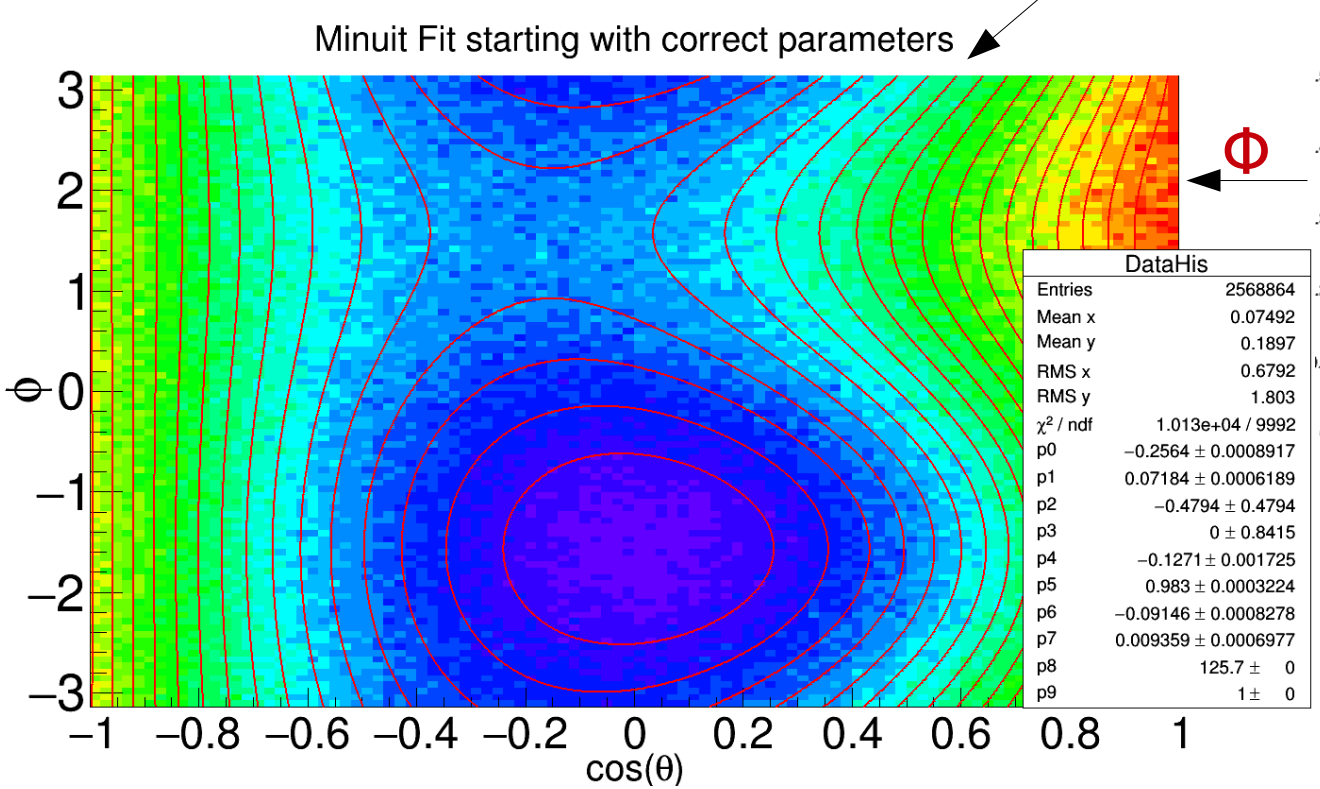
Helicity Amplitudes and partial waves

$$g(z) = \frac{1}{k}\Sigma_L\left[(L+1)T_L^+ + LT_L^-\right]P_L(z)$$

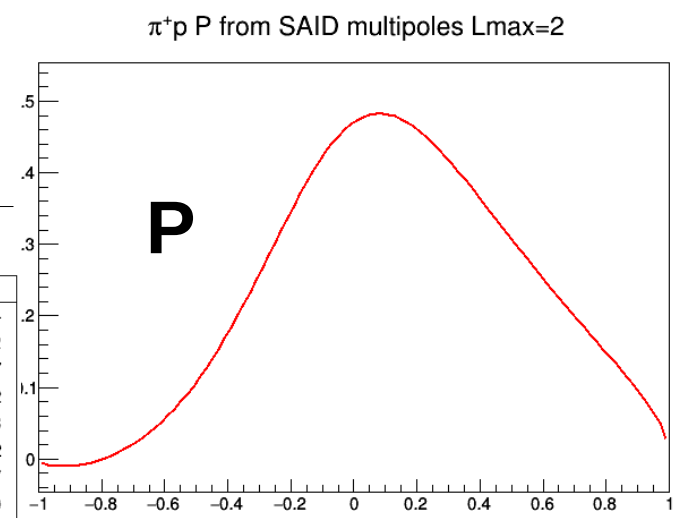
$$h(z) = \frac{1}{k}\Sigma_L\left[T_L^+ - T_L^-\right]\sqrt{1-z^2}P_L'(z)$$



cos(θ)



Φ



cos(θ)

MultiNest PW Fit

Truncate $L_{max}=1$

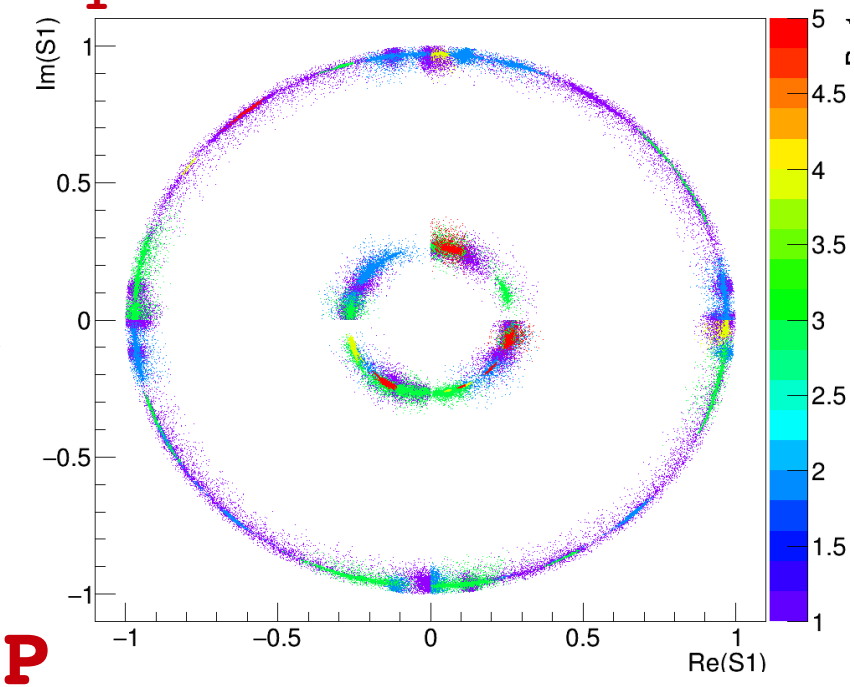
PW : S_1, P_1, P_3

Colours indicate different "solutions"

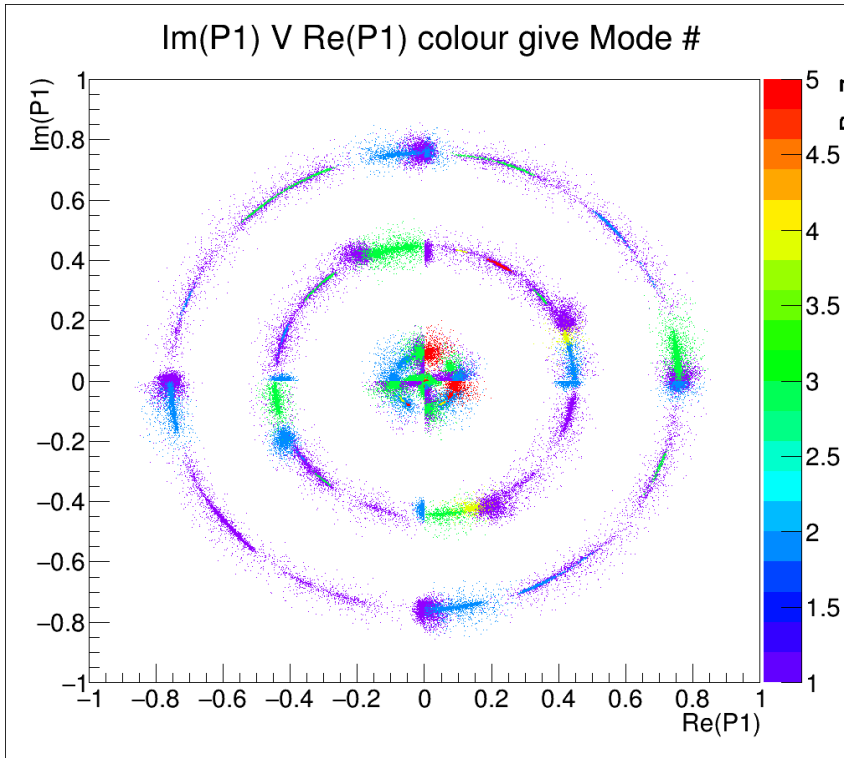
~10,000 live points

~ 2 hours

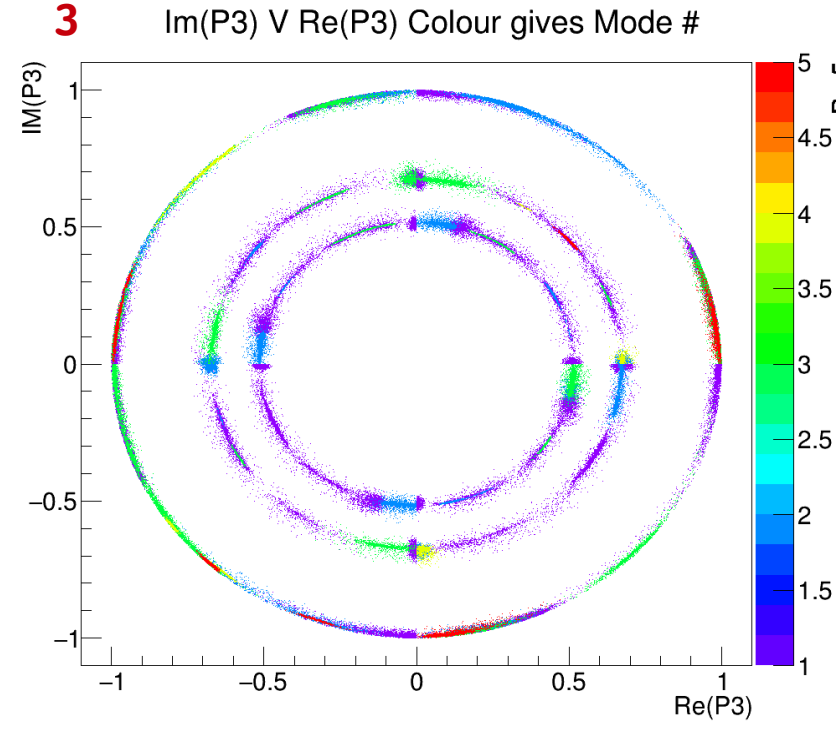
S_1



P_1

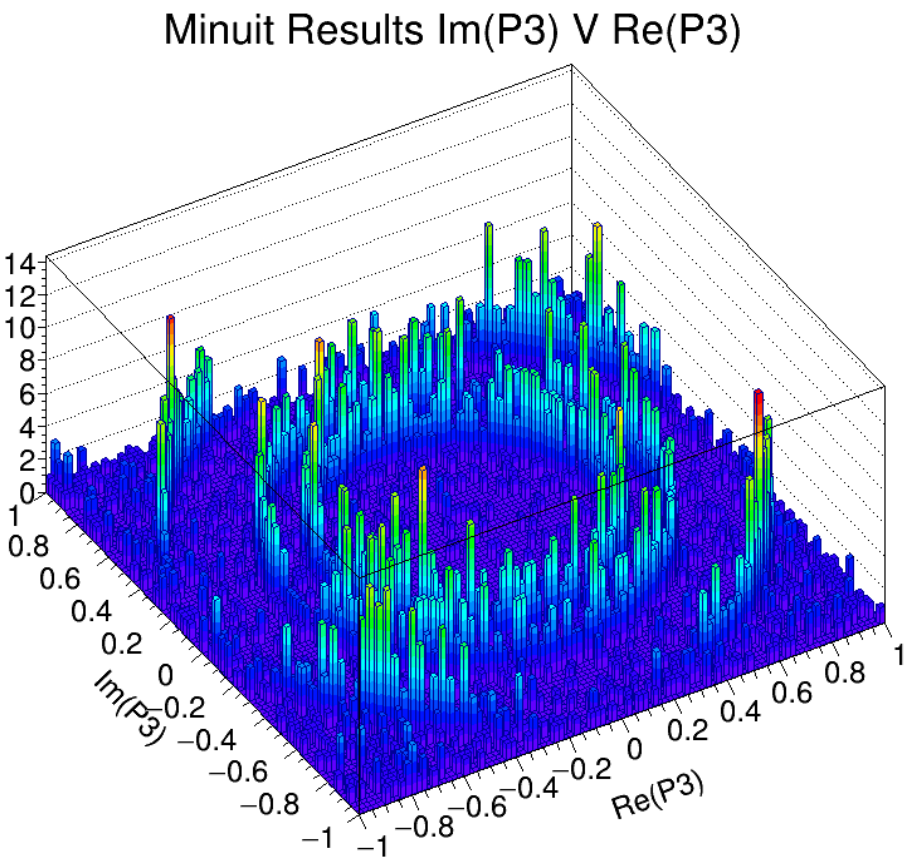


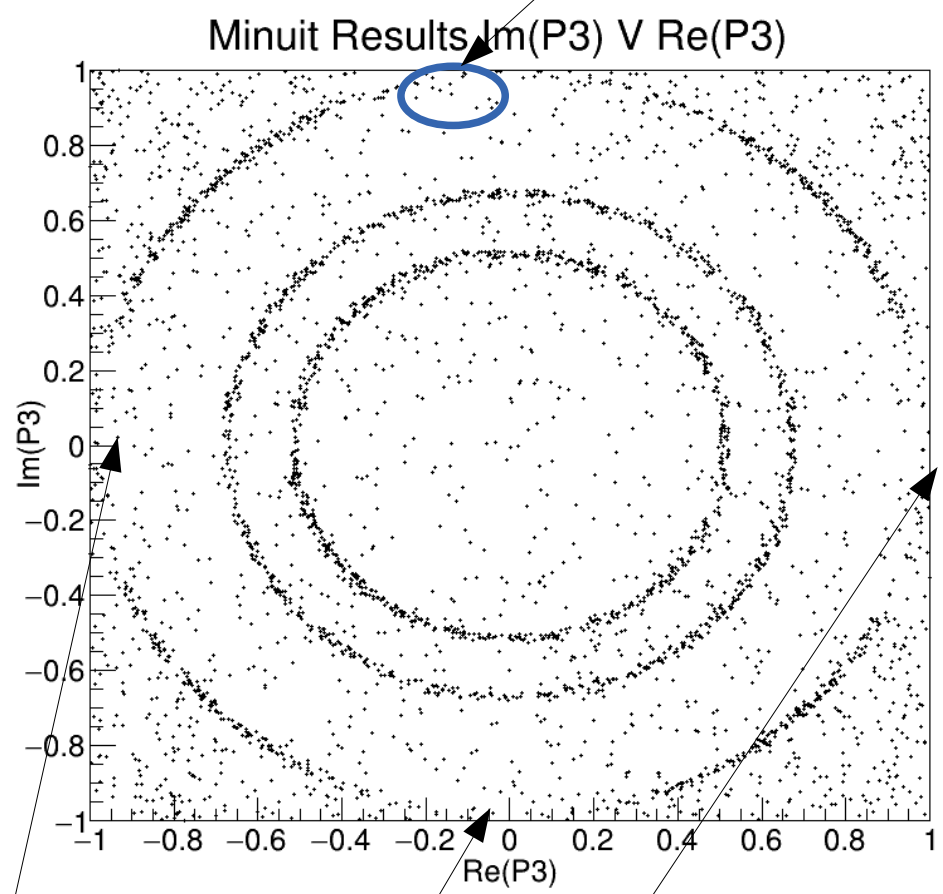
P_3



Use Minuit with Random initial values

Physical Value

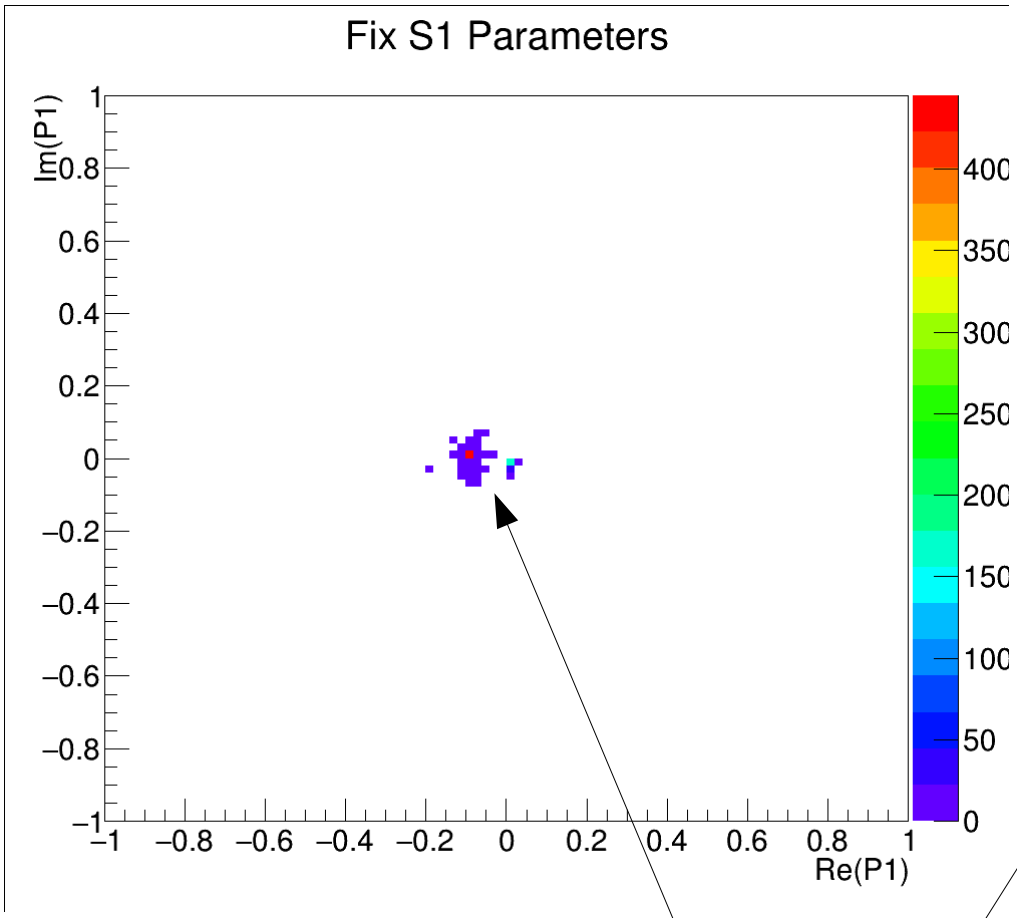


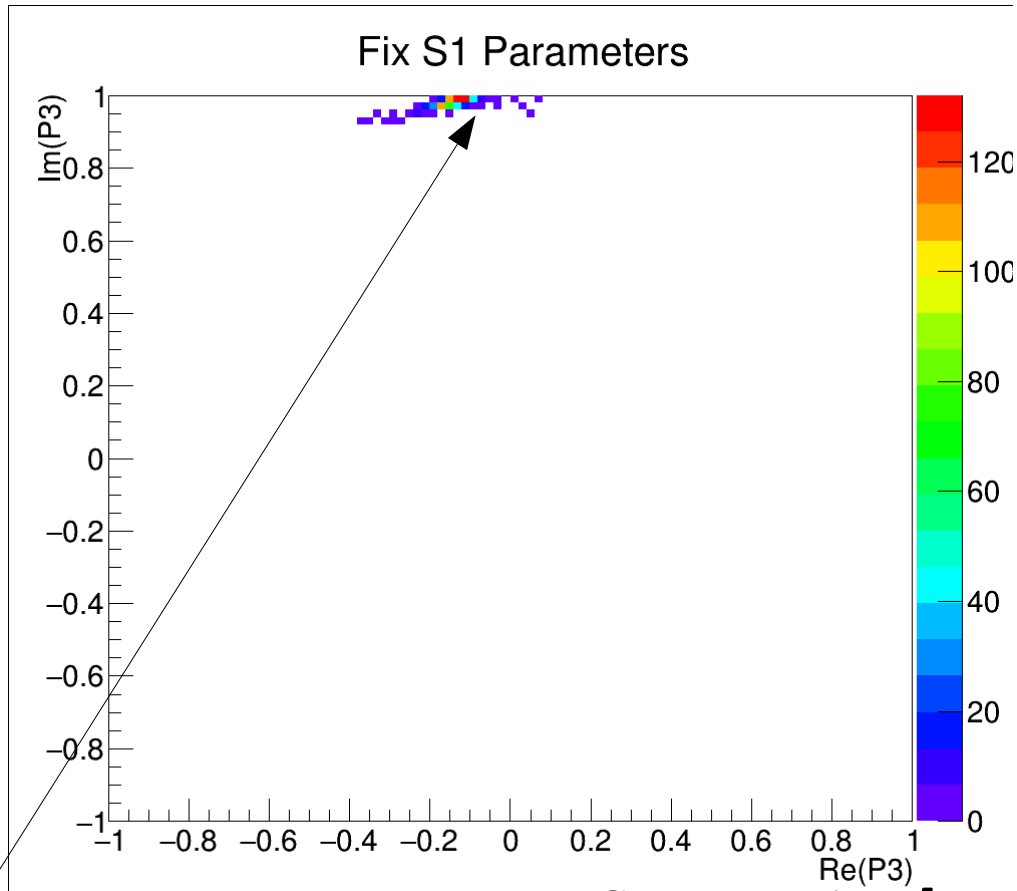


Lack of solutions where parameter close to 1

4000 fits ~ 3.5hours (1s a fit)
~ 2M likelihood calculations

Fix S1 parameters to true value

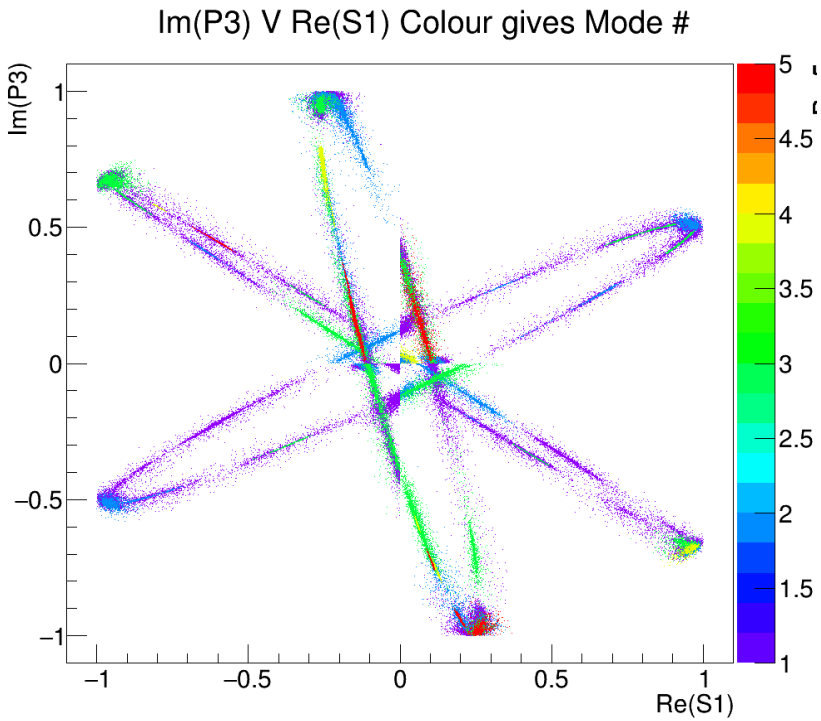


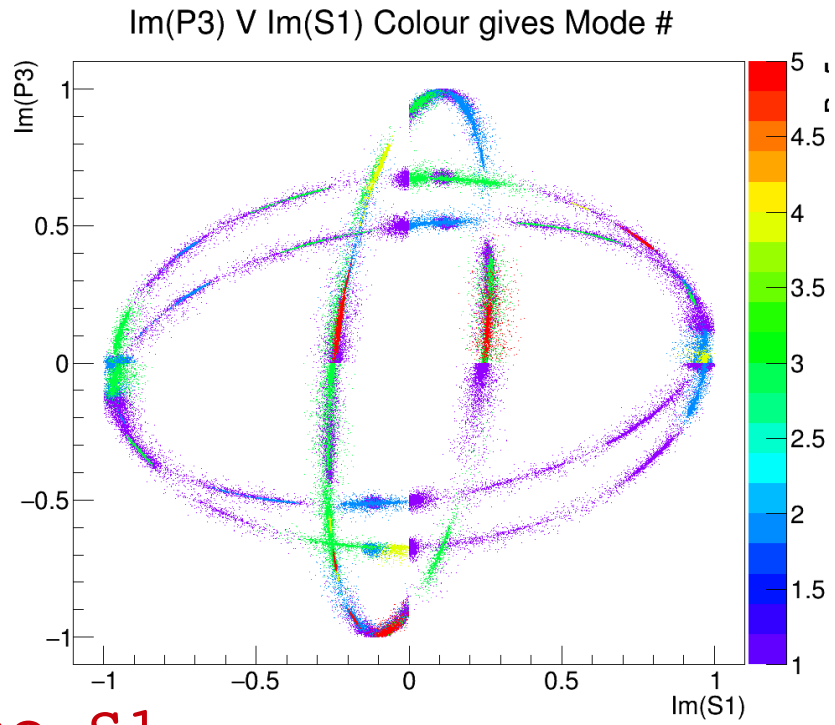


MultiNest finds:
Physical solution

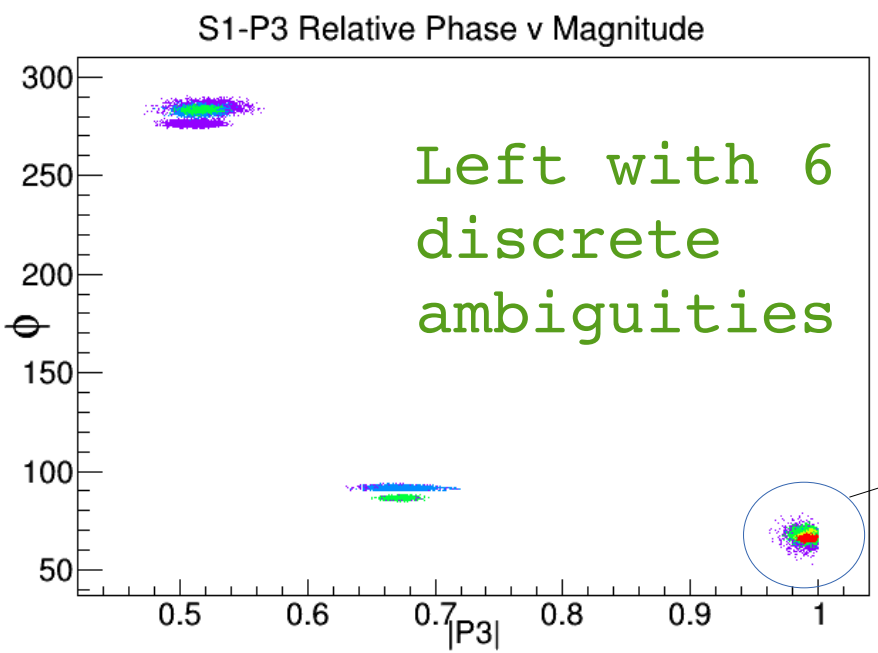
		Generated
S1	-0.256 ± 0.003	-0.257
	0.072 ± 0.010	0.072
P3	-0.123 ± 0.038	-0.126
	0.983 ± 0.005	0.983
P1	-0.091 ± 0.001	-0.091
	0.010 ± 0.004	0.009

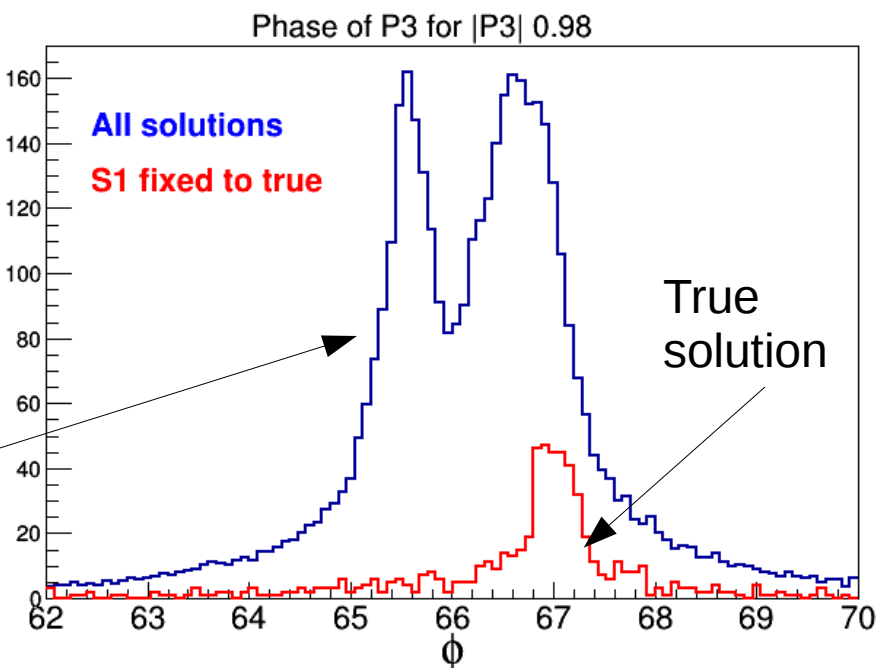
Relative Partial Waves





Look at P3 phase relative to S1





Summary

CLAS12 experiment will soon produce mesons through Quasi-real photoproduction

Currently preparing analysis framework to
- Handle large statistics datasets
- Analyse many final states
- Provide alternative methods

Investigated different signal/background separation

Investigated effectiveness of Longitudinal P.S.

Implemented Nested Sampling algorithm into AmpTools
- Investigating its usefulness in Amplitude Analysis

Currently implementng amplitudes in collaboration with JPAC and testing on available CLAS data

Developing Amplitude Analysis

CLAS : HASPECT MESONEX Light Meson Decay

Joint Physics Analysis Centre JPAC

Unitarity, Analyticity,
Crossing Symmetry,
Low/High Energy Constraints
Veneziano Amplitudes
Van Hove Plots
...

GluEx COMPASS BES Others...

In general greater overlap between different experiments and theorists

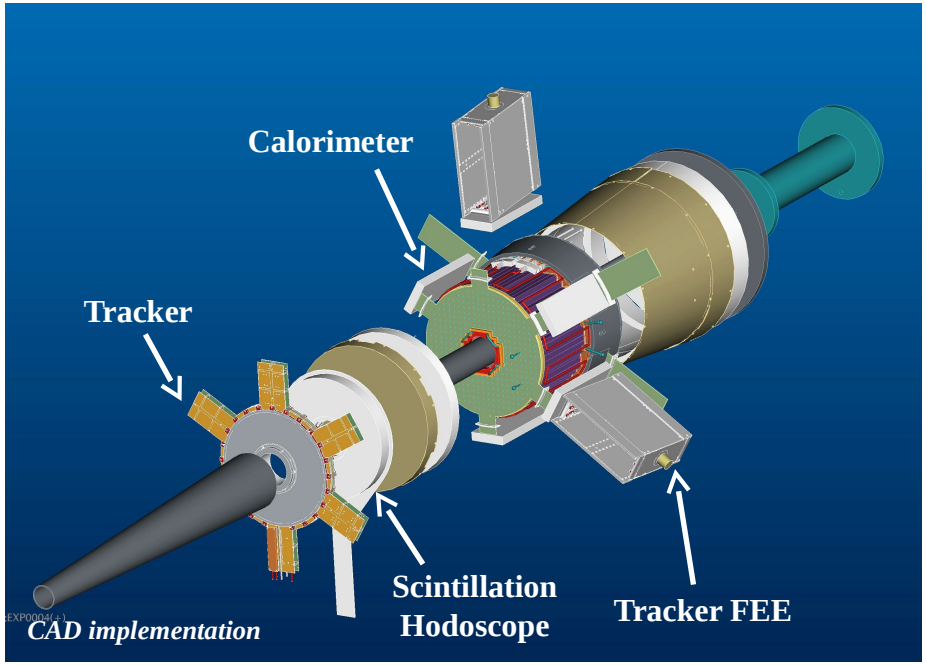
CLAS12 - Forward Tagger

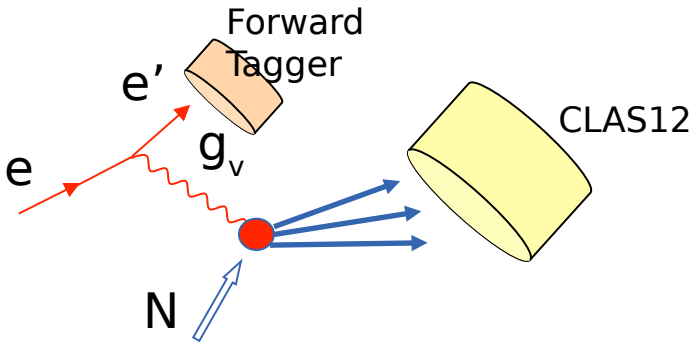
Detect electrons at small angle to perform quasi-real photo-production experiments.

Calorimeter: electron energy/momentum
Photon energy (ν=E-E')
Polarization $\varepsilon^{-1} \approx 1 + \nu^2/2EE'$
$PbWO_4$ crystals with APD/SiPM readout

Scintillation Hodoscope: veto for photons
Scintillator tiles with WLS readout

Tracker: electron angles, polarization plane
MicroMegas detectors





$E_{scattered}$	0.5 - 4.5 GeV
θ	2.5^o - 4.5^o
ϕ	0° - 360°
ν	6.5 - 10.5 GeV
Q^2	0.01 - 0.3 GeV^2 ($< Q^2 >$ 0.1 GeV^2)
W	3.6 - 4.5 GeV

CLAS12 Detector Systems

Forward Detector

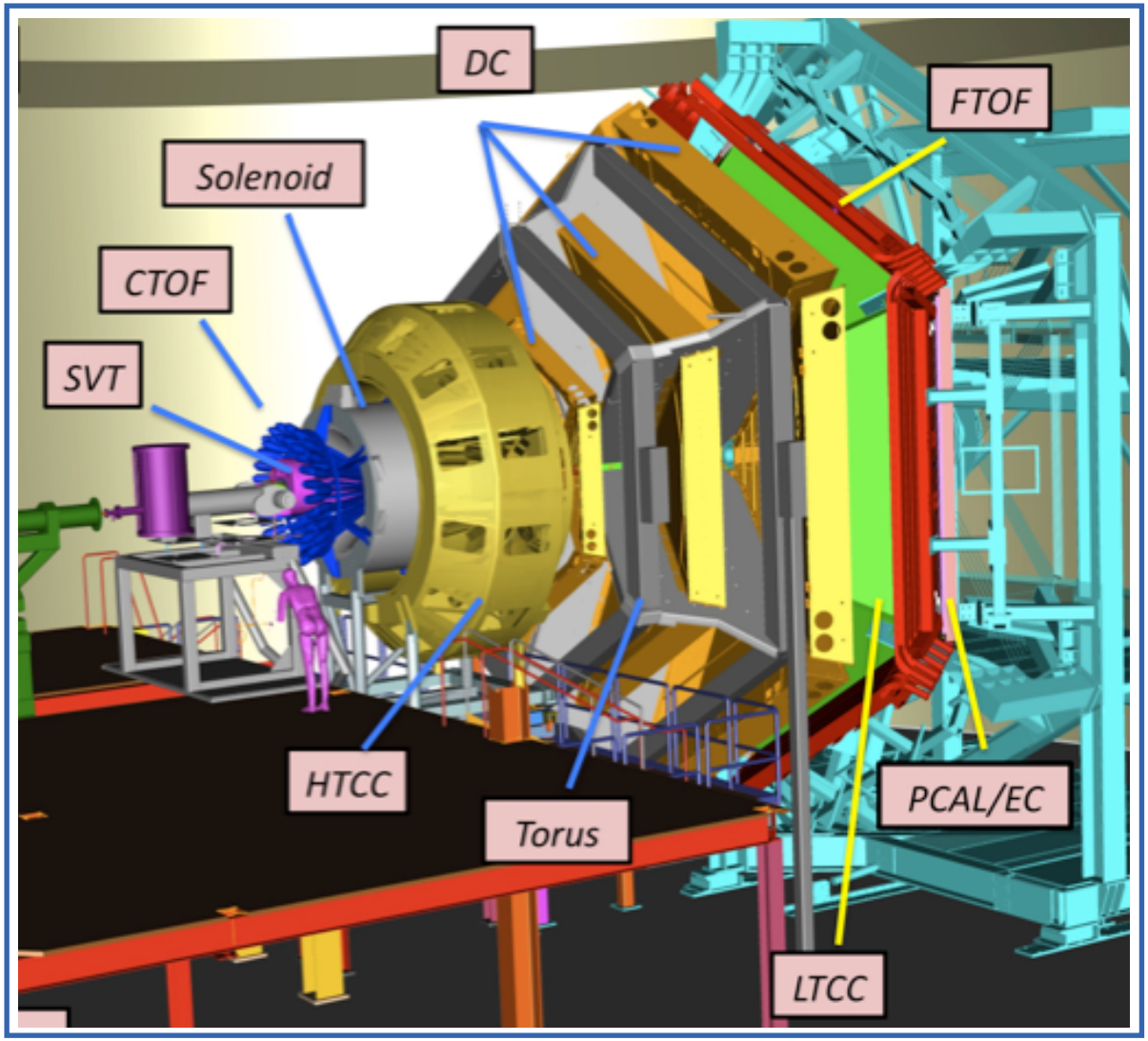
- TORUS Magnet
- Forward silicon vertex tracker
- HThresh Cerenkov Counter
- LThresh Cerenkov Counter
- Forward TOF System
- Preshower calorimeter
- E.M. Calorimeter

Central Detector

- SOLENOID magnet
- Barrel silicon tracker
- Central TOF

Additional Equipment

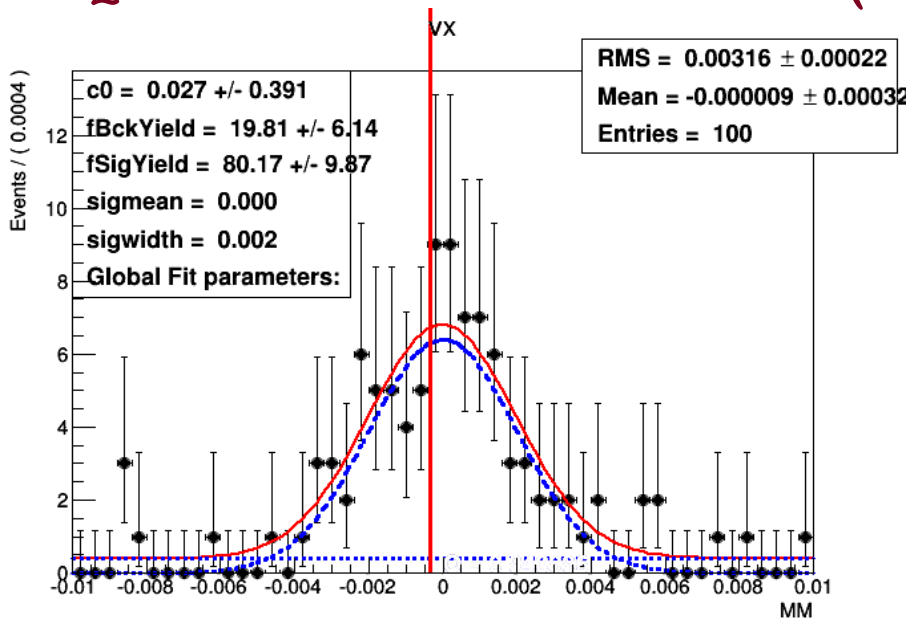
- Micromegas (CD)
- Neutron detector (CD)
- Forward RICH
- Forward Tagger

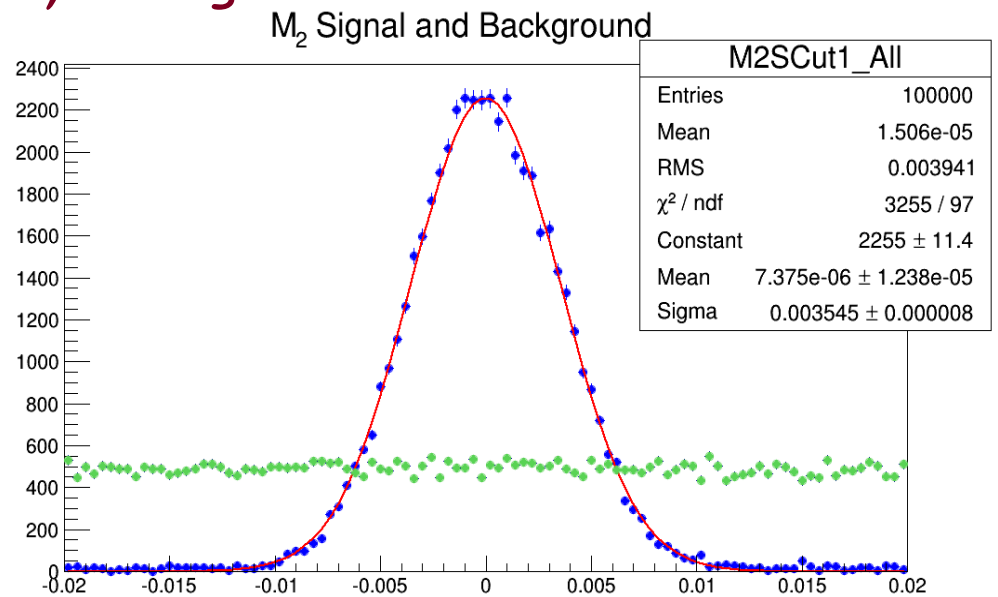


Enable e- detection below 5°

Ready for data Summer 2017

Qvalue with fixed (true) signal width





Qvalue reproduces signal and background shape

But uncertainties not correct (need to calculate)

And if width not constrained ...